Optimierung von Zugprioritäten in der Eisenbahnbetriebsdisposition mit Hilfe aktiver und deduktiver Datenbanken

Von der Fakultät für Bauingenieurwesen der
Rheinisch-Westfälischen Technischen Hochschule Aachen
zur Erlangung des akademischen Grades
eines Doktors der Naturwissenschaften
genehmigte Dissertation

vorgelegt von

Christian Sakowitz

Berichter: Universitätsprofessor Dr.-Ing. Ekkehard Wendler
Universitätsprofessor Dr.rer.nat. Rainer Manthey

Tag der mündlichen Prüfung: 31. März 2011

Diese Dissertation ist auf den Internetseiten der Hochschulbibliothek online verfügbar.

Inhaltsverzeichnis

Optimierung von Zugprioritäten in der Eisenbahnbetriebsdisposition mit Hilfe aktiver und deduktiver Datenbanken

– Zusammenfassung –

Die Maximierung der Gewinne ist ein elementares Ziel jedes wirtschaftlich agierenden Eisenbahninfrastrukturbetreibers. Um dieses Geschäftsziel im Zusammenhang mit der Eisenbahnbetriebsdisposition zu erreichen, müssen Verspätungen vermieden und fahrplanmäßige Anschlüsse sichergestellt werden. Allerdings sind in den meisten Fällen nominell gleich große Verspätungen zweier unterschiedlicher Züge nicht unbedingt unter wirtschaftlichen Gesichtspunkten gleichwertig. Ferner können bestimmte Anschlüsse zwischen Zügen unter diesem Aspekt bedeutender als andere sein.

Im Eisenbahnverkehr bestehen komplexe Abhängigkeiten und Wechselwirkungen. Unter Berücksichtigung dieser Effekte muß der Betriebsdisponent verschiedene Konfliktlösungsmöglichkeiten und daraus resultierende Wartezeiten überprüfen und sich dann für die bestmögliche Lösung entscheiden. Dies ist bei einem kurzfristig auftretenden zeitkritischen Konflikt allerdings praktisch unmöglich. Geschlossene mathematische Optimierungsverfahren stoßen in größeren Eisenbahnnetzen wegen der hohen Anzahl der dafür benötigten Nebenbedingungen rasch an ihre Grenzen.

In dieser Arbeit wird daher vorgeschlagen, den eigentlichen Optimierungsprozeß vom Betriebsdispositionsprozeß abzutrennen. Dabei sollen mit Hilfe aktiver, deduktiver und normativer Regeln unter wirtschaftlichen Gesichtspunkten ermittelte Zugprioritäten für die Konfliktlösung verwendet werden. Vorhandene „intelligente“ Datenbankverwaltungssysteme (DBMS) mit integrierten aktiven und deduktiven Datenbankfunktionalitäten können für diese Anwendung verwendet werden. Ein aktives DBMS ermöglicht die Definition von automatischen Systemreaktionen des DBMS auf bestimmte detektierte Datenbankereignisse. Ein deduktives DBMS erlaubt es, neue erschließbare Fakten zu spezifizieren, zu verwalten und aus expliziten Basisfakten abzuleiten.

Die Ermittlung und Vergabe von Zugprioritäten erfolgt separat vom Tagesgeschäft der Disposition des leicht bzw. schwer gestörten Eisenbahnbetriebs. Die langfristige Optimierung dieser Prioritäten erfolgt auf Basis der Auswertung von protokollierten Betriebsdaten.

Optimising train priorities for the regulation of train services using active and deductive databases

– Summary –

Profit maximisation is a fundamental goal of any business-driven railway infrastructure company. In order to achieve this target in the context of traffic regulation, it must try to avoid delays and ensure scheduled connections. However, nominally equal delays to two different trains are not equal in value from an economic point of view in most cases. Moreover, some connections between trains might be more important in this sense than others.

There are complex interdependencies and reciprocal effects in railway traffic. Considering these effects, a dispatcher must evaluate possible forms of conflict resolution and the waiting times these give rise to and select the best solution possible. This is not achievable where a time-critical conflict arises at short notice. Even closed mathematical optimisation algorithms encounter limits in the case of larger railway networks due to the enormous number of constraints to be considered.

This thesis will therefore propose that the optimisation process be separated from the train regulation process. Instead, economically evaluated train priorities for conflict situations are to be determined with the help of active, deductive and normative rules. Existing concepts of “smart” database management systems (DBMS) with integrated active and deductive database functionalities can be used for this application. An active DBMS allows the definition of reactions to be automatically initiated by the DBMS in response to the detection of given database-related events. A deductive DBMS allows new, deducible facts to be specified, administered and specially derived from explicitly introduced facts.

Train priorities are generated and assigned in detachment from day-to-day operations for lightly and heavily disrupted railway traffic respectively. Long-term optimisation of these priorities is effected by evaluating past operational data.

1 Einleitung

Eisenbahnbetrieb kann aufgrund plötzlich und unvorhersehbar eintretender Störungsereignisse nicht immer entsprechend dem vorgegebenen Fahrplan durchgeführt werden. Aufgabe der Eisenbahnbetriebsdisposition ist es daher, im Falle von Störungen unverzüglich geeignete Maßnahmen zu ergreifen, die eine schnellstmögliche Rückkehr zum planmäßigen Verkehr ermöglichen. Dies hat zwei zentrale Aspekte: die Minimierung der Verspätungsminuten für die Züge sowie die Reduzierung von Fahrgast- bzw. Gütertransportverspätungsminuten für die gesamte Reise- bzw. Transportkette. Dabei resultieren die Störungen von Fahrzeugen und Fahrwegelementen in konkreten Konflikten zwischen Zügen, die vom Disponenten durch spezielle Maßnahmen gelöst werden müssen (vgl. [Jac03]). Bei der Lösung dieser Maßnahmen spielen Prioritäten eine wichtige Rolle.

Diese Konflikte können in verschiedene Klassen eingeteilt werden [Bär96]. Zum einen gibt es Belegungskonflikte zwischen mindestens zwei Zügen. Aufgrund des Spurführungsprinzips der Eisenbahn und der technischen Sicherung können bestimmte, fest abgegrenzte Blockstrecken nur von maximal einem Zug belegt werden. Wenn nun mehrere Züge zum selben Zeitpunkt auf ein solches Belegungselement zugreifen möchten, ergibt sich zwischen diesen Zügen ein Konflikt, der auf Grundlage einer Prioritätenentscheidung zugunsten eines bestimmten Zuges aufgelöst werden muß. Für bzgl. der Priorität niederrangige Züge resultiert die Konfliktlösung in der Regel in einer Fahrtzeitverlängerung in der Größenordnung von mehreren Minuten. Als mögliche Konfliktlösungen sind zuungunsten des niederrangigen Zuges je nach konkreter Situation die Nutzung eines alternativen Fahrwegs, die Verlängerung eines planmäßigen Halts, die Verlegung von Überholungshalten, die Einfügung eines zusätzlichen Betriebshaltes oder eine Fahrtzeitverlängerung möglich.

Zum anderen können speziell in größeren Knotenbahnhöfen Anschlußkonflikte zwischen Zügen auftreten. Anschlüsse dienen dem Übergang von Reisenden bzw. Wagengruppen von einem Zubringerzug zu einem oder mehreren Anschlußzügen. Insbesondere im Personenverkehr ist bei einer Verspätung des Zubringerzuges in der Größenordnung von wenigen Minuten dieser Übergang unter Umständen ohne Verschiebung der Abfahrtszeit (und damit Fahrtzeitverlängerung) des Anschlußzuges nicht mehr möglich. Im Güterverkehr sind die Toleranzen in der Regel etwas größer. Bei einem Anschlußkonflikt muß eine Prioritätenentscheidung entweder zugunsten der Reisenden bzw. des Transportgutes des Zubringerzuges (Anschluß erhalten) oder zugunsten der Pünktlichkeit des Anschlußzuges (Anschluß aufgeben) erfolgen. Bei einer Entscheidung zuungunsten des Anschlußzuges können für diesen Fahrtzeitverlängerungen im Bereich von mehreren Minuten auftreten. Bei einer Entscheidung zuungunsten des Zubringerzuges können für die gesamten Reiseketten der umsteigewilligen Reisenden Verzögerungen bis in den Stundenbereich auftreten. Beim Gütertransport können sich bei Übergängen von Wagengruppen in einem Rangierbahnhof sogar Verzögerungen von mehreren Stunden ergeben.

Einen Spezialfall des Anschlußkonfliktes stellt der Umlaufkonflikt dar. Hierbei geht es um Anschlüsse eines Zuges „auf sich selbst“, z. B. bei Wendezügen am Endbahnhof einer Linie. Die spezielle Eigenschaft dieser Anschlüsse ist die Tatsache, daß ein Brechen des An-

schlusses offensichtlich aus technischen Gründen nicht möglich ist. Eine Möglichkeit zur Konfliktlösung besteht ggf. darin, den Wendevorgang durch Auslassung einzelner Halte auf einen vorgelagerten Bahnhof zu verlegen, so daß der Anschlußzug wieder entsprechend der ursprünglichen Planung die Fahrt aufnehmen kann. Auch hier ist wieder eine Prioritätenentscheidung zwischen der pünktlichen Fahrt des Anschlußzuges und dem Bedienen aller vorgesehenen Verkehrshalte erforderlich. Hier stehen wiederum alternativ Fahrtverlängerungen zweier Beförderungsgruppen gegenüber.

Man kann schließlich bei der Disposition noch den Spezialfall eines Gleissperrkonfliktes betrachten. Dieser Konflikt ist eine Abwandlung des Belegungskonfliktes für einen Zug, bei dem ein Belegungselement wegen externer Ursachen (z. B. Wartungsarbeiten) über einen bestimmten Zeitraum nicht angefordert werden kann. Auf solch einen Konflikt muß wie auf einen Belegungskonflikt reagiert werden, wobei der Zug per Definition als gegenüber dem externen Ereignis niederrangig betrachtet werden muß.

Um diese prioritätenbasierten Konfliktlösungen optimieren zu können, ist ein geeigneter Maßstab erforderlich. Als solche können durch die europäische Gesetzgebung [EG07, EG01] vorangetriebene und in vielen Bereichen bereits implementierte Bonus-Malus-Systeme dienen. Mit Hilfe eines solchen Systems kann die verkehrswirtschaftliche Gewichtung von Pünktlichkeit und Anschlußsicherheit für einen Zug quantifiziert werden. Damit kann für einen Verkehrshalt eines Zuges eine Bewertungsfunktion der Pünktlichkeit in Abhängigkeit von den Verspätungsminuten aufgestellt werden. Ebenso können auch Pönale für nicht eingehaltene Anschlüsse festgelegt werden. Auf Basis des Maßstabes dieses Bonus-Malus-Systems sollten nun die Prioritätenentscheidungen so festgelegt werden, daß die daraus resultierenden Konfliktlösungen zu möglichst geringen Bewertungen führen.

Im aktuellen Stand von Wissenschaft und Technik (Abb. 1-1) (vgl. [Hel04]) ist eine automatische Konflikterkennung und -lösung prinzipiell möglich. Hier gibt es verschiedene Ansätze, die von geschlossenen mathematischen Optimierungsverfahren bis hin zu heuristischen Verfahren reichen. Im Rahmen der heuristischen Verfahren, genauer gesagt, der Simulationsverfahren werden heute auch schon Prioritäten zur Optimierung der Konfliktlösungen verwendet. Diese Prioritäten sind derzeit aber in der Regel statisch und werden manuell vorgegeben. Damit kann die schwankende verkehrswirtschaftliche Bedeutung einer Zugfahrt nur unzureichend wiedergegeben werden. Bonus-Malus-Systeme sind bei Eisenbahnen erst seit kurzer Zeit in Kraft und werden daher natürlich auch noch nicht unmittelbar bei der Optimierung der Betriebsdisposition verwendet.

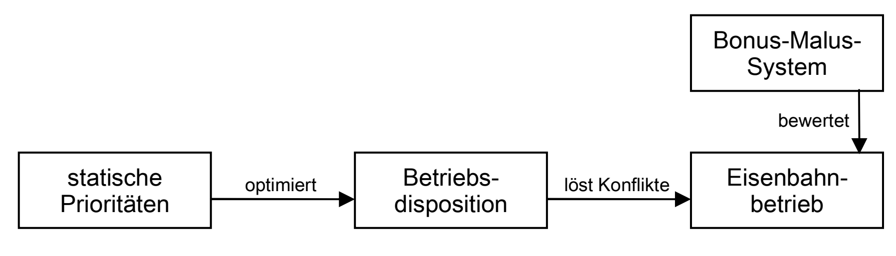


Abb. 1-1: Aktueller Stand der Technik

Im Rahmen dieser Arbeit wird über den aktuellen Stand von Wissenschaft und Technik hinausgehend vorgeschlagen, anstelle von statischen Prioritäten zukünftig dynamische Prioritäten zu verwenden und in einer Weise kontinuierlich automatisch anzupassen, daß eine Dispositionsoptimierung auf Basis eines vorgegebenen Bonus-Malus-Systems erfolgen kann (Abb. 1-2). Ideal wäre in diesem Falle eine geschlossene Optimierung, z. B. mit Hilfe der linearen Optimierung. Dies läßt sich praktisch aber nur bei relativ dünn befahrenen Netzen mit relativ einfacher Topologie in Echtzeit realisieren. Ansonsten steigt entweder der erforderliche Rechenaufwand wegen der großen Anzahl der erforderlichen Nebenbedingungen zu stark an oder die Nebenbedingungen werden dermaßen vereinfacht, daß das wirkliche Optimum unter Umständen nicht mehr gefunden wird.

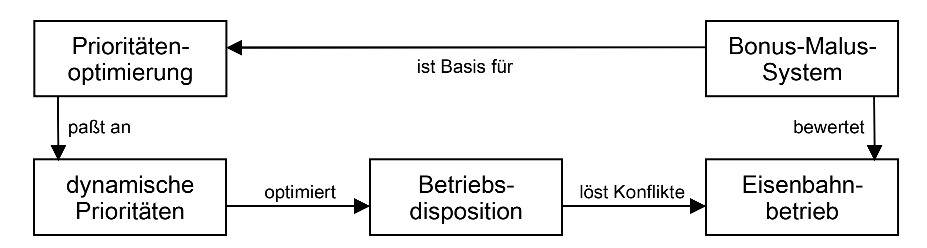


Abb. 1-2: Vorschlag für ein neues Verfahren

Daher wird ein neuartiges System zur Konfliktlösung vorgeschlagen, das die Vorteile der heuristischen Verfahren (geringe Rechenzeit) mit den Vorteilen der geschlossenen Optimierungen (Güte der Ergebnisse bzgl. der verkehrswirtschaftlichen Bewertung entsprechend des vorgegebenen Bonus-Malus-Systems) verbindet. Dazu werden die Prozesse prioritätenbasierte Betriebsdisposition, Bewertung der Dispositionsergebnisse anhand des Bonus-Malus-Systems und die eigentliche Prioritätenoptimierung in unterschiedlichen Teilsystemen (Dispositionssystem, Qualitätsüberwachungssystem und Prioritätenvergabe- und -optimierungssystem) voneinander entkoppelt. Die Prioritätenanpassung erfolgt außerhalb des eigentlichen zeitkritischen Tagesgeschäftes der Konflikterkennung und -lösung durch den Disponenten in einem eigenständigen Prioritätenvergabe- und -optimierungssystem. Dieses soll

- die bei einer prioritätenbasierten Disposition zu verwendenden Zug- und Anschlußprioritäten verwalten,
- eine nach klar nachvollziehbaren Kriterien automatisch erfolgende Prioritätenvergabe im Rahmen der Disposition ermöglichen,
- anstelle von statischen Prioritäten dynamische Prioritäten für die Disposition generieren können,
- die Folgen von Störfälle in der Disposition durch entsprechende Anpassung von Dispositionsprioritäten berücksichtigen und
- ein Optimierung der Dispositionsergebnisse durch eine kontinuierliche Anpassung der intern verwalteten Prioritäten ermöglichen

Durch die Anwendung des hier vorgeschlagenen Verfahrens resultieren viele Vorteile. Je nach Implementierung des Bonus-Malus-Systems kann der wirtschaftliche Erlös von Eisenbahnverkehrs- und -infrastrukturbetreibern erhöht werden. Durch die europaweite gesetzli-

che Verankerung von Bonus-Malus-Systemen ergeben sich langfristig geeignete Rahmenbedingungen zur Integration des Ansatzes in die „Lebenswirklichkeit“ der Eisenbahnen. Der Disponent wird von der bisher erforderlichen, politisch wegen hohem Diskriminierungspotential brisanten manuellen Vergabe statischer Prioritäten entlastet. Der Infrastrukturverwalter ist im Falle von Diskriminierungsvorwürfen in der Lage, Prioritätenentscheidungen eindeutig auf Datenbasis zu begründen. Die Disposition wird durch die Verwendung von dynamischen Prioritäten „flexibler“, d. h. es können in jedem Einzelfall verstärkt die durch das Bonus-Malus-System vorgegebenen lokalen verkehrswirtschaftlichen Parameter berücksichtigt werden. Ferner kann durch ein Prioritätenvergabe- und -optimierungssystem eine kurzfristige (bzgl. der Bewertung der Disposition sinnvolle) automatische Reaktion auf bestimmte Störfälle im Sinne einer Prioritätenanpassung durchgeführt werden. Schließlich ist eine automatische langfristige Evaluierung des mit den Prioritäten disponierten Eisenbahnbetriebs anhand des Bonus-Malus-Systems möglich. Darauf kann mit kontinuierlicher Überprüfung und ggf. Anpassung der verwendeten Dispositionsprioritäten reagiert werden.

Für die technische Umsetzung des Prioritätenvergabe- und -optimierungssystems wird in dieser Arbeit anstelle der traditionell bei eisenbahnbetrieblichern Systemen durchgeführten Implementierung mittels einer objektorientierten, imperativen oder deklarativen Programmiersprache eine regelbasierte Umsetzung unter Verwendung vorhandener Konzepte der Datenbanktechnologie vorgeschlagen (vgl. z. B. [Kem97]). Im Zentrum des Interesses stehen dabei deduktive, aktive und normative Komponenten der Datenbankverwaltung (DBMS). Mit einer deduktiven Komponente des DBMS können neuer Fakten aus wenigen expliziten Basisfakten und einer Regelmenge abgeleitet werden. Eine aktive Komponente erlaubt die Definition automatisch ausgeführter Systemreaktionen als Reaktion auf die Erkennung bestimmter datenbankrelevanter Ereignisse. Wiederum erfolgt die Spezifikation dieser Reaktionen in einem Regelformat. Mit einer normativen Komponente ist es möglich, feste Integritätsbedingungen zu definierten, die von der Datenbank in jedem konsistenten Zustand eingehalten werden müssen.

Die deduktive Komponente kann insbesondere im Zusammenhang mit der Definition von für die Prioritätenoptimierung relevanten Ereignissen (Auftreten ungünstiger Dispositionsergebnisse) verwendet werden. Die aktive Komponente kann unter anderem zur automatischen Ausführung von Reaktionen (Prioritätenanpassungen) eingesetzt werden. Die normative Komponente dient der Festlegung illegaler Prioritätenkonstellationen, z. B. aufgrund gesetzlicher oder betrieblicher Vorgaben.

Vorteile eines datenbankbasierten Implementierungsansatzes liegen vor allem darin begründet, daß der Implementierungs- und langfristige Wartungsaufwand deutlich reduziert werden kann. Entscheidend ist in diesem Zusammenhang die Trennung zwischen elementarer Daten- und Ereignisverarbeitung und der eigentlichen Systemlogik. Die Details zur internen Datenorganisation, -zugriff, -manipulation und -integrität sowie zur kontinuierlichen Ereigniserkennung und -signalisierung sind für den Systementwickler nicht sichtbar und werden vom bereitgestellten DBMS eigenständig verantwortet. Der Entwickler kann sich deshalb darauf beschränken, den mittels einer Modellierungssprache vorab erstellten Systementwurf entsprechend in den vom DBMS unterstützten Regelsprachen umzusetzen und

diese Regeln langfristig z. B. aufgrund geänderter gesetzlicher Rahmenbedingungen anzupassen. Diese Regeln sind im Gegensatz zu traditionellem Programmcode häufig selbstdokumentierend, so daß die dahinterstehende Systemlogik offensichtlich erkennbar bleibt. Aufgrund dieser Vorgehensweise ist die Prioritätenvergabe als solches für die davon betroffenen Gruppen (insbesondere Eisenbahnverkehrsunternehmen und Regulierungsbehörden) transparenter darstellbar, besser dokumentierbar und leichter adaptierbar.

Regelbasierte Systeme werden gemäß [Lau10] im beschränkten Umfang bereits in eisenbahnbetrieblichen Anwendungen erfolgreich eingesetzt oder sind Gegenstand aktueller Forschung, hier allerdings nicht im Datenbankkontext sondern im Zusammenhang mit semantisch verwandten Expertensystemen in einem objektorientierten Programmierumfeld. Zu den vorhandenen und untersuchten Anwendungen gehören Simulationen, Fahrwegsicherungen in elektronischen Stellwerken, Datenvalidierungen und Signalmodellierungen. Auf Anwendungen im Zusammenhang mit Eisenbahnbetriebsdisposition wird ausführlich in Kapitel 2 eingegangen.

Die Grundidee des hier vorgestellten Verfahrens wurde vom Verfasser dieser Arbeit bereits in [Sak06] veröffentlicht.

Die Arbeit hat den folgenden Aufbau: Zunächst werden in Kapitel 2 die Grundlagen zur prioritätengesteuerten Eisenbahnbetriebsdisposition ausführlich dargestellt, woraufhin sich in Kapitel 3 eine Vorstellung der verwendeten Modellierungs- und Spezifikationstechniken anschließt. Auf Basis dieser Grundlagen erfolgt in Kapitel 4 die Modellierung der Anwendungsumgebung eines Prioritätenvergabe- und -optimierungssystems. In Kapitel 5 schließt sich die Darstellung des vorgeschlagenen Verfahrens zur Prioritätenvergabe und -optimierung an, auf dessen Basis in Kapitel 6 ein detaillierter regelbasierter Systementwurf erstellt wird. Schließlich erfolgt in Kapitel 7 eine Vorstellung des im Rahmen dieser Arbeit prototypisch implementierten Prioritätenvergabe- und -optimierungssystems POS.

2 Prioritätengesteuerte Eisenbahnbetriebsdisposition – Stand der Wissenschaft

In diesem Kapitel wird die im aktuellen Stand der Wissenschaft und Technik praktizierte prioritätengesteuerte Eisenbahnbetriebsdisposition detailliert vorgestellt. Nach einer kurzen Einführung in die Grundlagen des Eisenbahnbetriebs erfolgt eine Vertiefung auf Qualitätsmaßstäbe. Anschließend werden die Verfahren zur Betriebsdisposition allgemein ausgeführt, woraufhin sich eine Fokussierung auf prioritätenbasierte Verfahren anschließt. Schließlich werden die Defizite erläutert.

2.1 Grundlagen des Eisenbahnbetriebs

2.1.1 Übersicht

Im folgenden sollen einige wesentliche, im Rahmen der Arbeit behandelte und zum weiteren Verständnis erforderliche Grundbausteine und Fachbegriffe des Eisenbahnbetriebs eingeführt werden. Für eine umfassende wissenschaftliche Einführung in alle Facetten des Eisenbahnbetriebs sei auf die entsprechende Fachliteratur (z. B. [Pac04]) verwiesen. Die Darstellung konzentriert sich auf den Eisenbahnbetrieb in der Bundesrepublik Deutschland. Die Unterschiede zu den übrigen, insbesondere europäischen Staaten sind in der Regel technischer, aber nicht grundsätzlicher Natur.

Das Anwendungsgebiet Eisenbahnbetrieb – soweit es im Rahmen dieser Arbeit von Interesse ist – kann in vier zentrale, nicht disjunkte Sichten (Abb. 2-1) unterteilt werden:

- normative Betriebssicht
- technische Betriebssicht
- Planbetriebssicht
- operative Betriebssicht

Die normative Betriebssicht beinhaltet die vom Gesetzgeber, der Exekutive und dem Betreiber festgelegten Gesetze, Verordnungen und Richtlinien für den Eisenbahnbetrieb. Damit werden „normative Regeln“ für den Eisenbahnbetrieb festgesetzt. Die technische Sicht enthält die „Hardware“ des Eisenbahnbetriebs, nämlich den Fahrweg und die Fahrzeuge, sowie das technisch realisierte Betriebsverfahren. Die Planbetriebssicht beschreibt den im Rahmen des Eisenbahnbetriebs umzusetzenden und vorab konstruierten Fahrplan. Die operative Betriebssicht enthält bestimmte im Rahmen des operativen Eisenbahnbetriebs auftretende besondere Ereignisse, z. B. Störungen und Planabweichungen.

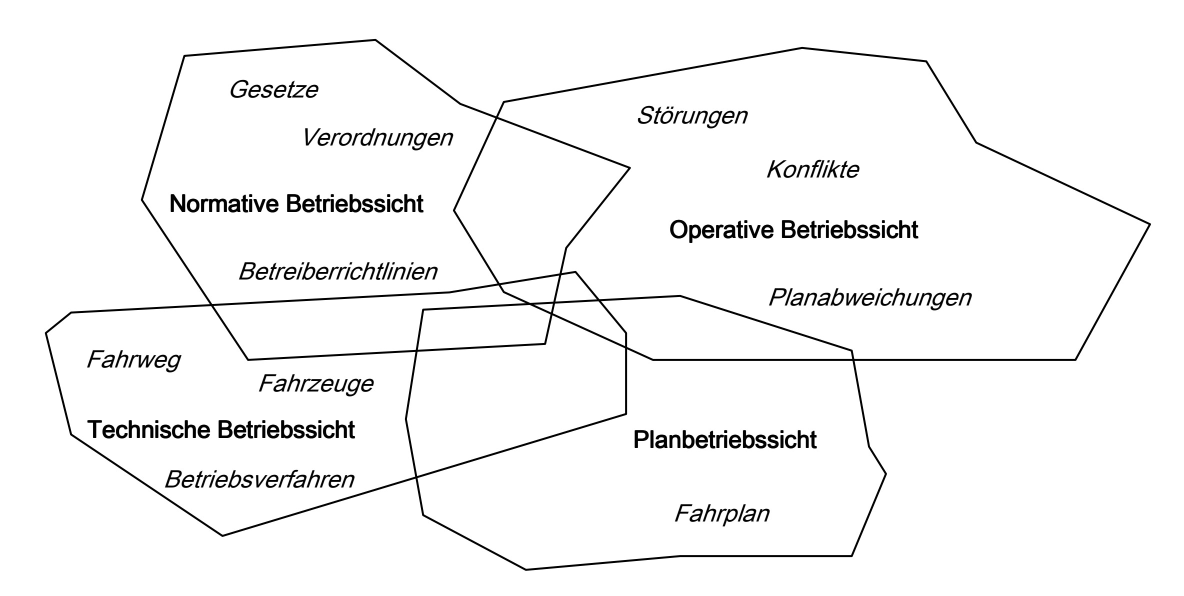


Abb. 2-1: Sichten des Anwendungsgebietes Eisenbahnbetrieb

2.1.2 Normative Betriebssicht

Das Eisenbahnrecht ist ein umfangreiches Rechtsgebiet. Eine ausführliche Darstellung ist z. B. [Kun05] zu entnehmen. Die Systematik der normativen Vorschriften kann Abb. 2-2 entnommen werden.

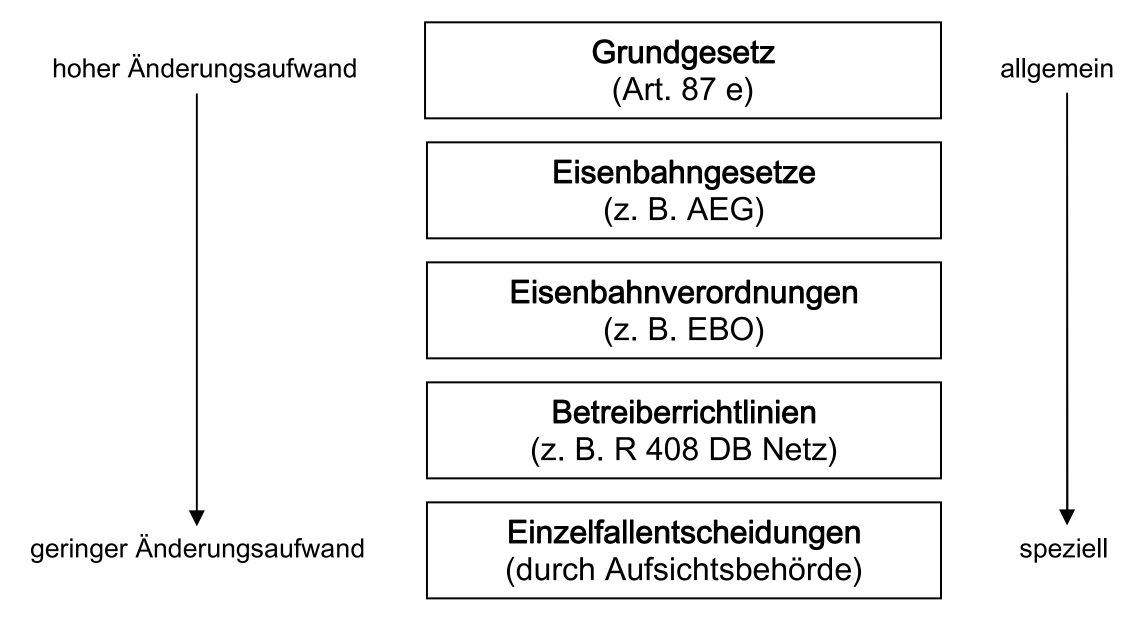


Abb. 2-2: Systematik normativer Vorschriften in Deutschland

Neben den verfassungsrechtlichen Vorgaben über die Verwaltung von Eisenbahnen des Bundes in Artikel 87 e des Grundgesetzes [GG49] gibt es weitere gesetzliche Vorschriften, die den Eisenbahnbetrieb betreffen. Man unterscheidet hierbei zwischen zwei Arten von Schienenbahnen: der Bereich der Straßen-, Stadt- und U-Bahnen wird durch das Personenbeförderungsgesetz (PBefG) [PBefG90] geregelt, für die in dieser Arbeit betrachteten Haupt- und Nebenbahnen sind Landeseisenbahngesetze oder das Allgemeine Eisenbahngesetz

(AEG) [AEG93] gültig. Das AEG wird seit Beginn der 1990er Jahre auch stark durch das Recht der Europäischen Gemeinschaft bzw. Europäischen Union beeinflußt. Als Beispiel sei die Richtlinie 91/440/EWG [EWG91] genannt, die den Prozeß der Bahnreform durch eine stärkere Trennung von Eisenbahninfrastrukturbetreibern und Verkehrsanbietern und den damit verbundenen offenen und diskriminierungsfreien Netzzugang in die Wege geleitet hat. Unterhalb der Gesetzesebene befindet sich die Ebene der Verordnungen, unter denen hier stellvertretend die Eisenbahn-Bau- und Betriebsordnung (EBO) [EBO67] und die Eisenbahn-infrastruktur-Benutzungsverordnung (EIBV) [EIBV05] genannt werden sollen. Unterhalb der Verordnungen sind noch spezielle regulierende Richtlinien der Infrastrukturbetreiber anzusiedeln. Für den Bereich der DB Netz wäre hier insbesondere die Richtlinie 408 „Züge fahren und Rangieren“ [DBR408] zu erwähnen. Daneben sind Einzelfallentscheidungen der jeweiligen Aufsichts- und Genehmigungsbehörden zu berücksichtigen.

Auf im Rahmen dieser Arbeit besonders zu berücksichtigende Vorschriften wird direkt im jeweiligen Zusammenhang eingegangen. Es soll an dieser Stelle aber noch darauf hingewiesen werden, daß die normative Betriebssicht (insbesondere in den letzten Jahren durch den Prozeß der Bahnreform) fortwährend starken inhaltlichen Änderungen ausgesetzt ist und man hier schwer mittel- und längerfristige Prognosen über zukünftige Entwicklungen machen kann.

2.1.3 Technische Betriebssicht

Die technische Betriebssicht umfaßt zum einen den Fahrweg, zum anderen die auf dieser Eisenbahninfrastruktur verkehrenden Fahrzeuge. Schließlich wird noch das technisch von Fahrweg und Fahrzeug realisierte Betriebsverfahren erläutert. Eine ausführliche Darstellung der technischen Aspekte kann z. B. [Pac04] entnommen werden.

Der **Fahrweg**, der durch einen Eisenbahninfrastrukturbetreiber bewirtschaftet wird, besteht aus ein- oder mehrgleisigen Strecken. Diese Strecken werden in Knotenbahnhöfen und Abzweigstellen miteinander verknüpft. Auf den Strecken befinden sich Bahnhöfe und Haltepunkte, die als Zugangsstellen zur Eisenbahn dienen. In Bahnhöfen und auf Strecken können die Bahnhofs- bzw. Streckengleise durch Weichen miteinander verbunden sein.

Die Eisenbahn-Bau- und Betriebsordnung (§ 4) [EBO67] definiert in diesem Zusammenhang wichtige Begriffe und Zusammenhänge:

- *Bahnanlagen:*
  alle Grundstücke, Bauwerke und sonstigen Einrichtungen einer Eisenbahn, die unter Berücksichtigung der örtlichen Verhältnisse zur Abwicklung oder Sicherung des Reise- oder Güterverkehr auf der Schiene erforderlich sind. Dazu gehören unter anderem:
  - *Bahnhöfe:*
    Bahnanlagen mit mindestens einer Weiche, wo Züge beginnen, enden, ausweichen oder wenden dürfen

  - *Blockstrecken:*
    Gleisabschnitte, in die ein Zug nur einfahren darf, wenn sie frei von Fahrzeugen sind
  - *Blockstellen:*
    Bahnanlagen, die eine Blockstrecke begrenzen. Eine Blockstelle kann zugleich als Bahnhof, Abzweigstelle, Überleitstelle oder Haltepunkt eingerichtet sein
  - *Abzweigstellen:*
    Blockstellen der freien Strecke, wo Züge von einer Strecke auf eine andere Stelle übergehen können
  - *Überleitstellen:*
    Blockstellen der freien Strecke, wo Züge auf ein anderes Gleis derselben Strecke übergehen können
  - *Haltepunkte:*
    Bahnanlagen ohne Weichen, wo Züge planmäßig halten, beginnen oder enden dürfen
  - *Hauptgleise:*
    von Zügen planmäßig befahrenen Gleise
  - *Durchgehende Hauptgleise:*
    Hauptgleise der freien Strecke und ihre Fortsetzung in den Bahnhöfen
  - *Nebengleise:*
    sonstige Gleise

Als Oberbegriff für Bahnhöfe, Haltepunkte und Abzweigstellen wird in dieser Arbeit der Begriff „Betriebsstelle“ verwendet (siehe [Brü95]).

Bedingt durch die Verwendung aktiver Weichen und eine fahrwegseitige Sicherungstechnik hat der Fahrweg bei der Durchführung des Eisenbahnbetriebs die zentrale Schlüsselstellung inne.

Im Zusammenhang mit der Durchführung des Eisenbahnbetriebs spielen technische Gesichtspunkte der **Fahrzeuge** eher eine untergeordnete Rolle gegenüber der Fahrwegssteuerungs- und -sicherungstechnik, so daß hier nur einige wenige Aspekte genannt werden sollen.

Die Fahrzeuge werden jeweils von einem Eisenbahnverkehrsunternehmen betrieben. Sie können auf dem Fahrweg eines oder mehrerer Eisenbahninfrastrukturbetreiber verkehren.

Die Eisenbahn-Bau- und Betriebsordnung (§§ 18-24) [EBO67] definiert wiederum:

- *Regelfahrzeuge:*
  alle Triebfahrzeuge und Wagen (sonstige Fahrzeuge spielen im Rahmen dieser Arbeit keine Rolle)
  - *Triebfahrzeuge:*
    werden unterteilt in Lokomotiven, Triebwagen und Kleinlokomotiven. Sie wer-

den entweder unmittelbar bedient oder werden gesteuert. Steuerung ist die Regelung der Antriebs- und Bremskraft durch eine Steuereinrichtung von einem führenden Fahrzeug aus oder durch Fernsteuerung.

- *Wagen:*
  werden eingeteilt in Reisezugwagen und Güterwagen

Der Aspekt der Steuerung durch Regelung der Antriebs- und Bremskraft ist der zentrale Beitrag der Fahrzeuge zum Eisenbahnbetrieb.

Als spezielle Eigenart des Verkehrsträgers Schiene gibt es beim Eisenbahnbetrieb ein starkes technisches Zusammenspiel zwischen dem Fahrweg und den darauf verkehrenden Fahrzeugen. So werden die Weichen fahrwegseitig gesteuert, wohingegen das Beschleunigen und Bremsen fahrzeugseitig realisiert werden. Da der Bremsweg des Fahrzeuges jedoch den Sichtweg des Triebfahrzeugführers bei den heutzutage auf der Schiene gefahrenen Geschwindigkeiten deutlich überschreitet, obliegt der Fahrwegseite die Sicherstellung einer freien Strecke. Daher ist in Deutschland durch EBO [EBO67] und für den Bereich der DB Netz zusätzlich durch die Richtlinie 408 [DBNe04] (sowie in den meisten anderen Staaten ebenfalls) als **Betriebsverfahren** das „Fahren im Raumabstand" vorgegeben. Dazu wird der Fahrweg in Blockstrecken unterteilt, in denen sich zu jedem Zeitpunkt höchstens ein Zug befinden darf. Die entsprechende Blockstrecke darf seitens eines für die Betriebssicherheit verantwortlichen Stellwerks erst wieder für einen anderen Zug zur Einfahrt freigegeben werden, wenn der vorherige Zug die Blockstrecke zuzüglich eines sogenannten Durchrutschweges hinter der Blockstrecke (die Zugschlußstelle) vollständig passiert hat. Ansonsten muß der Zug am zugehörigen Blocksignal warten, bis die entsprechende Bedingung erfüllt ist. Die Einhaltung muß durch eine aufwendige Sicherungstechnik am Fahrweg und im Fahrzeug und zusätzliche Rückfallebenen im Stellwerk streng überwacht werden. Um eine von der Sicherungstechnik unbeeinflußte Fahrt zu gewährleisten, muß die Blockstrecke sogar bereits frei sein, wenn die Zugspitze den Bremseinsatzpunkt (BEP) für ein rechtzeitiges Abbremsen vor Erreichen der Blockstelle passiert, um ein rechtzeitiges Anhalten gewährleisten zu können. Bedingt durch die Sicherungstechnik (Schalt- und Reaktionszeiten) erhöht sich die sogenannte Belegungs- bzw. Sperrzeit eines Zuges für die unbehinderte Fahrt durch eine Blockstrecke noch um einen bestimmten von der Art der Sicherung abhängigen Zeitzuschlag.

2.1.4 Planbetriebssicht

Der Begriff „Fahrplan" ist nicht allgemeingültig definiert. Er legt den Fahrtverlauf von Zügen vorausschauend fest (siehe [Pac04]). Unter einem „Zug" wird in diesem Zusammenhang allerdings nicht primär ein technisches Fahrzeug verstanden, sondern ein spezielles Produktangebot eines Eisenbahnverkehrsunternehmens, das in diesem Fahrplan genauer spezifiziert bzw. beworben wird. Wesentliche Angaben für Züge sind eine nicht unbedingt eindeutige Produktbezeichnung (in der Regel Kombination aus einer Produktklasse und Zugnummer), die Verkehrstage, ein Laufweg, Ankunfts-, Abfahr- und Durchfahrzeiten in allen Betriebsstellen und zulässige Geschwindigkeiten in den einzelnen Abschnitten des Laufwegs. Vor allem im Güterverkehr und vereinzelt im Personenverkehr kann der Zug außer an

seinem Ausgangs- und Bestimmungsort in weiteren Betriebsstellen auf seinem Laufweg einzelne Wagengruppen aufnehmen bzw. abgeben. Zusätzlich können im Fahrplan Anschlüsse vorgegeben werden, bei denen Reisenden (durch individuellen Umstieg oder Wagengruppenübergang) bzw. Frachtgut (durch Wagengruppenübergang) in einem Bahnhof ein Wechsel von einem Zug auf einen anderen Zug ermöglicht wird.

Durch Aufstellung eines Fahrplans für den Eisenbahnbetrieb sollen mehrere Ziele erreicht werden: Einmal dient er der Koordination der Züge mit sich möglicherweise wegen überschneidenden Belegungszeiten gegenseitig ausschließenden Wunschfahrlagen. Ferner ermöglicht er die Information der (potentiellen) Kunden. Außerdem beschreibt der Fahrplan den gewünschten Betriebsablauf und ist damit eine wichtige Grundlage des Eisenbahnbetriebs. Schließlich kann der Fahrplan zur Untersuchung der Leistungsfähigkeit und Erkennung von Engpaßabschnitten der Eisenbahninfrastruktur verwendet werden (siehe dazu [Wen99]).

Eine wichtige Anforderung an den Fahrplan ist, daß dieser „wahr“ ist, also wie geplant auch tatsächlich im Betrieb durchgeführt werden kann [Gro66]. Dies ist der Fall, wenn die vorgegebenen Fahrzeiten unter realen Betriebsbedingungen einzuhalten sind und sich bei pünktlichem Betrieb keine Behinderungen zwischen den Zügen ergeben. Die Fahrzeiten können dabei mit Methoden der Fahrdynamik aus dem Zusammenspiel von Zugkräften und fahrzeug- und fahrwegbezogenen Widerständen bestimmt werden (siehe hierzu [Wen83]).

Um die Behinderungsfreiheit sicherzustellen, kann auf die bereits in 2.1.3 erwähnten Sperrzeiten für die Blockstrecken zurückgegriffen werden, die den Zeitbedarf für eine unbehinderte Durchfahrt gewährleisten. Der Fahrplan ist genau dann behinderungsfrei, wenn sich dort zu keinem Zeitpunkt in einer Blockstrecke Sperrzeiten von Zügen teilweise überschneiden. (Genaues Verfahren in [Hap59] bzw. [Brü95].) Das Sperrzeitenmodell wurde seitens des Eisenbahnweltverbandes UIC in der Richtlinie 406 [UIC04] zum Standardmodell für Streckenkapazitätsbetrachtungen erklärt.

Zur weiteren Vertiefung der Planbetriebssicht wird z. B. auf [Grö02] verwiesen.

2.1.5 Operative Betriebssicht

Der operative, tatsächlich durchgeführte Betrieb kann in größeren Netzen meistens nicht exakt nach Fahrplan durchgeführt werden. Es treten aufgrund von externen, technischen und betrieblichen Ursachen gelegentlich temporäre Störungen auf, die den Fahrweg, die Fahrzeuge oder das Personal betreffen können. Die Störungen resultieren in aller Regel in operativen Fahrplanabweichungen der davon direkt betroffenen Züge, sogenannten Urverspätungen. In deren Folge können Belegungskonflikte (vgl. [Pfe95]) mit anderen Zügen auftreten, d. h. ein pünktlicher langsamer (bzw. verspäteter schneller) Zug erreicht den Bremseinsatzpunkt vor einer noch durch einen schnelleren, verspäteten (bzw. langsameren, pünktlichen) Zug belegten Blockstrecke und wird dadurch abgebremst. Durch diese betrieblichen Behinderungen können weitere Planabweichungen (Folgeverspätungen) entstehen. Aus diesen Folgeverspätungen können dann durch weitere Belegungskonflikte wieder neue Fol-

geverspätungen resultieren usw. Eine langfristige Sperrung eines Streckenabschnittes kann auch als eine besondere Form eines Belegungskonfliktes interpretiert werden.

In [Ack98] wird vorgeschlagen, auch Gleisverlegungen innerhalb einer Station als eigenständige Planabweichungen zu betrachten. Da Gleisverlegungen bei frühzeitiger Disposition und dem Vorhandensein geeigneter Informationssysteme für abfahrende Fahrgäste unproblematisch sind, und für ankommende Fahrgäste ohnehin nur im Zusammenhang mit bereits betrachteten Anschlüssen von Relevanz sind, sollen diese in der weiteren Darstellung als eigenes Qualitätskriterium des Eisenbahnbetriebs vernachlässigt werden.

Verspätungen führen für die Kunden der Eisenbahnverkehrsunternehmen zu unmittelbaren Beförderungs- bzw. Transportzeitverlängerungen. Außerdem besteht die Möglichkeit, daß planmäßig vorgegebene Anschlüsse auf andere Züge eventuell nicht mehr gewährleistet werden können (Anschlußkonflikt [Bär96]). Dieses Nichteinhalten von Anschlüssen kann neben den Verspätungen als zusätzliche Planabweichung angesehen werden, die nochmals zu teilweise erheblichen Reise- bzw. Transportzeitverlängerungen führen können. Schließlich kann es auch zu Störungen des Umlaufs von Fahrzeugen und Personal kommen (Umlaufkonflikt [Bär96]). Daraus resultieren ggf. weitere Planabweichungen in Form von Verspätungen weiterer Züge und mögliche Anschlußverluste.

Zur weiteren Vertiefung der operativen Betriebssicht wird z. B. auf [Hel04] verwiesen.

2.2 Qualität des Eisenbahnbetriebs

2.2.1 Übersicht

Planabweichungen sind aus der Sicht der Endkunden der Eisenbahn, also der Fahrgäste bzw. Verlader, als qualitätsmindernd zu betrachten. Bei zu großen betrieblichen Qualitätsmängeln besteht die Gefahr einer Verschiebung des sogenannten Modal-Splits, d. h. einer Kundenabwanderung auf konkurrierende Verkehrsträger. Genauere verkehrswirtschaftliche Zusammenhänge können [Wal91] und [Joc99] entnommen werden.

Wegen der großen Bedeutung der Betriebsqualität ist es sinnvoll, ein entsprechend geeignetes „Gütemaß“ für den Eisenbahnbetrieb zu definieren. Damit ist es möglich, die Qualität verschiedener Eisenbahnbetriebsabläufe in vorgegebenen Untersuchungszeiträumen miteinander zu vergleichen.

Für die Definition eines einheitlichen Gütemaßes haben sich in letzter Zeit aufgrund verschiedener gesetzlicher Initiativen, insbesondere der Europäischen Kommission, Bonus-Malus-Systeme als zunehmend implementierte Lösungen herausgestellt. Mittels eines solchen Bonus-Malus-Systems kann ein genereller, einheitlicher und flexibler Ansatz zur Aufstellung von Gütemaßen verfolgt werden, indem der Gegenwert von Pünktlichkeit und Anschlußsicherheit für einen Zug ökonomisch objektiv quantifiziert wird. Bonus-Malus-Systeme können sowohl auf dem Primärmarkt zwischen Endkunden und Verkehrsbetrieb als auch auf

dem Sekundärmarkt zwischen Verkehrsbetrieb und Infrastrukturbetreiber vereinbart werden, wobei eine weitestgehende Adäquatheit der Systeme angestrebt werden sollte.

Hinter einem Bonus-Malus-System steckt der Gedanke, daß auf dem Primärmarkt das Verkehrsunternehmen dem Endkunden für die Durchführung der Beförderung bzw. auf dem Sekundärmarkt der Netzbetreiber dem die Zugfahrt durchführenden Verkehrsunternehmen für Planung und betriebliche Abwicklung einen bestimmten Geldbetrag als „Basispreis" in Rechnung stellt. Wenn dabei entsprechend vereinbarte Qualitätsvorgaben nicht erfüllt werden, ist der Verschulder verpflichtet, eine Maluszahlung zu entrichten, die der von der Qualitätseinbuße betroffenen Partei als Bonus zugeführt wird. Auf dem Primärmarkt kann zum Beispiel dem Verlader aufgrund verspäteter Bereitstellung des Transportgutes eine Maluszahlung angerechnet werden, wohingegen er bei verspäteter Zustellung des Wagens durch das Verkehrsunternehmen einen Bonus erhält. Auf dem Sekundärmarkt ist das Verkehrsunternehmen bei einem defekten Triebfahrzeug zu einer Maluszahlung gegenüber dem Infrastrukturbetreiber verpflichtet, wohingegen es bei Verspätungen aufgrund von Verzögerungen im Betriebsablauf einen Bonus erhält.

Mit Hilfe von Bonus- bzw. Maluszahlungen kann man somit eine Erhöhung bzw. Erniedrigung des Basispreises relativ zur Betriebsqualität vornehmen. Damit soll einerseits der Gegenwert ausgedrückt werden, den der Endkunde bzw. das Eisenbahnverkehrsunternehmen für eine bestimmte Leistung zu zahlen bereit ist (verkehrswirtschaftlicher Wert eines Angebotes), andererseits soll der eigentliche Verursacher von vermeidbaren Betriebsstörungen pönalisiert werden. Somit sollen den beteiligten Unternehmen wirtschaftliche Anreize zur Qualitätssteigerung gegeben werden.

Diese Bonus-Malus-Zahlungen können in unterschiedlichen Verkehrssegmenten stark variieren. Züge des Personenverkehrs, die eine starke Zubringerfunktion für einen Taktknoten erfüllen, reagieren zum Beispiel deutlich empfindlicher auf Verspätungen als ein saisonaler Urlaubsreisezug. Im Güterbereich hängen die verkehrswirtschaftlichen Bewertungen sehr stark davon ab, inwieweit auf den Transport folgende Produktionsschritte durch eine Planabweichung gestört werden. Ein Extrembeispiel wäre ein Kraftwerk, das aufgrund deutlich verspäteter Kalklieferungen den Betrieb einstellen muß und dadurch enormen volkswirtschaftlichen Schaden verursachen kann.

Im folgenden sollen die gesetzlichen Grundlagen sowie bestehende Bonus-Malus-Systeme für den Eisenbahnbetrieb auf dem Primär- und Sekundärmarkt vorgestellt werden.

2.2.2 Bonus-Malus-Systeme auf dem Primärmarkt

Ein Bonus-Malus-System zwischen dem Endkunden und dem Eisenbahnverkehrsunternehmen ist im Personenverkehr gesetzlich durch den Artikel 17 der EU-Verordnung Nr. 1371/2007 [EG07] sowie § 17 der Eisenbahnverkehrsordnung [EVO38] festgelegt.

Hierbei erhalten Reisende 25 % das Fahrkartenwertes in bar zurückerstattet, wenn sie ihren Zielort wegen Verspätung bzw. Anschlußverlust mit mehr als 60 Minuten Verspätung erreichen, bei einer Verspätung von mehr als 120 Minuten werden 50 % erstattet. Dabei besteht

ein Mindesterstattungsbetrag von 4 €, so daß örtlicher Nahverkehr praktisch vor der Regelung ausgeschlossen ist. Ferner besteht im Falle einer Verspätung von mehr als 60 Minuten die Möglichkeit eines Reiserücktritts bei Erstattung des Fahrpreises. Falls das Reiseziel nicht mehr am selben Tage erreicht werden kann, stehen dem Reisenden eine Aufwandsentschädigung von bis zu 80 € für Taxi- bzw. Übernachtungskosten zu. Ausnahmen bestehen im Falle höherer Gewalt.

Im Güterverkehr werden Bonus-Malus-Systeme derzeit nicht gesetzlich vorgeschrieben. Man unterscheidet hier zwischen Ganzzug- und Einzelwagenverkehren. Bei Ganzzugverkehren gibt es für jeden Transport je nach Bedarf individuelle vertragliche Bonus-Malus-Regelungen zwischen Verkehrsunternehmen und Verladern. Für Abfahrtpünktlichkeit ist der Verlader verantwortlich, für Ankunftspünktlichkeit das Verkehrsunternehmen. Bei Einzelwagenverkehren, die in Deutschland nahezu exklusiv von der DB-Logistiktochter DB Schenker Rail und evtl. Unterauftragnehmern durchgeführt werden, wird zwischen den drei Produkten „Classic“, „Quality“ und „Prime“ unterschieden [Sti05]. Bei den letzten beiden (teureren) Produkten erstattet DB Schenker Rail für jeden unterhalb einer Erfüllungsquote von 95 % verspäteten Wagen 40 € bzw. 70 €. Ausnahmen bestehen, wenn die Verspätung durch den Absender (verzögerte Beladung) oder höhere Gewalt verursacht wurde. Güterverkehrsunternehmen im europäischen Ausland haben eigene Regelungen.

2.2.3 Bonus-Malus-Systeme auf dem Sekundärmarkt

Ein Sekundärmarkt zwischen Eisenbahnverkehrsunternehmen und Infrastrukturbetreibern entstand im europäischen Bereich erst in den letzten Jahren mit der durch die Richtlinie 91/440/EWG [EWG91] begonnenen Trennung von Eisenbahninfrastruktur und Verkehrsdienstleitungen inkl. der Anforderung an einen diskriminierungsfreien Netzzugang. Der Sekundärmarkt ist systembedingt im Regelfall ein monopolistischer „Markt“ mit einem in der Regel staatlichen Infrastrukturbetreiber, der von einer ebenfalls staatlichen Regulierungsbehörde überwacht wird.

Artikel 11 der EU-Richtlinie Nr. 2001/14 [EG01] sowie § 21 der Eisenbahninfrastruktur-Benutzungsverordnung [EIBV05] schreiben die Implementierung eines Bonus-Malus-Systems auf dem Sekundärmarkt als Bestandteile, die „den Eisenbahnverkehrsunternehmen und den Betreibern der Schienenwege Anreize zur Verringerung von Störungen und zur Erhöhung der Leistungsfähigkeit des Schienennetzes bieten“ explizit vor.

Details werden in dieser Richtlinie bzw. Verordnung allerdings offen gelassen, so daß hier ein großer Gestaltungsspielraum für die Infrastrukturbetreiber und Regulierungsbehörden verbleibt.

In Deutschland wurde bei der DB Netz AG zum 13.12.2009 das derzeit gültige „Anreizsystem“ eingeführt [DBNe09]. Hierzu werden für jede Verkehrsart (Fern-, Nah- und Güterverkehr) prozentuale Pünktlichkeitszielwerte festgesetzt, die über das Fahrplanjahr eingehalten werden sollen. Meßpunkt ist die Ankunftszeit am Zielbahnhof des Zuges, wobei ein Zug ab 6 Minuten (Personenverkehr) bzw. 31 Minuten (Güterverkehr) Planabweichung als verspätet

gilt. Wenn an einem Betriebstag das Pünktlichkeitsziel für eine Verkehrsart nicht erreicht wurde, werden vom Verkehrsunternehmen vorab ausgewählte Zugläufe dieser Verkehrsart (20 % der zugewiesenen Trassen) bzgl. verursachter Verspätungsminuten auf dem Zuglauf ausgewertet. Dabei werden ausschließlich Vorfälle berücksichtigt, die durch das Verkehrsunternehmen oder den Infrastrukturbetreiber direkt beeinflußbar sind. Folgeverspätungen werden ignoriert. Die verursachten Verspätungsminuten werden entsprechend dem Verursachungskonto des Verkehrsunternehmens bzw. von DB Netz angerechnet. Nur falls die Pünktlichkeitsziele über das gesamte Fahrplanjahr nicht eingehalten wurden, erfolgen Maluszahlungen entsprechend der über das Jahr erstellten Verspätungsminutenkonten. Die Pönalzahlung besteht aus 0,10 € pro erfaßter Verspätungsminute, wobei eine Obergrenze festgelegt ist.

Im Rahmen des von der Europäischen Kommission finanzierten Projektes New Opera [NeO08] wurden weitere europäische Bonus-Malus-Systeme auf dem Sekundärmarkt untersucht.

Im Vereinigten Königreich gilt folgendes Erstattungssystem zwischen Verkehrsbetrieben und dem Infrastrukturbetreiber NetworkRail: Bei einer Verspätung oberhalb von drei Minuten zahlt der Verursacher (Verkehrsunternehmen oder Infrastrukturbetreiber) dem davon betroffenen Verkehrsunternehmen pro Verspätungsminute einen Ausgleich. Die Abrechnung erfolgt dabei grundsätzlich über NetworkRail, da die konkurrierenden Verkehrsunternehmen keine direkten Vertragspartner sind. Im Falle von Unklarheit über den Verursacher einer Verspätung trifft der staatliche Eisenbahnregulator die Entscheidung. Zusätzlich wird jährlich eine Toleranzgrenze für Verspätungen festgelegt, bei Verletzung dieser Grenze durch den Infrastrukturbetreiber erhält das Verkehrsunternehmen einen Bonus.

Das „komplexeste" Bonus-Malus-System auf dem Sekundärmarkt ist in Italien eingeführt. Es basiert auf den Projekten EPR (European Performance Regime) I und II des Eisenbahnweltverbandes UIC bzw. der europäischen Infrastrukturbetreibervereinigung RNE. Nach einer Testlaufphase ist es seit dem 1.1.2005 im Einsatz. Es basiert ausschließlich auf Verspätungen und umfaßt alle auf dem Netz verkehrenden Züge. Ziel ist eine Qualitätsverbesserung jedes einzelnen Zuglaufs. Sowohl Verkehrsbetriebe als auch der staatliche Infrastrukturbetreiber RFI sind in gleicher Weise bei der Verursachung von Verspätungen zu analogen Erstattungsleistungen verpflichtet. Sämtliche Zahlungen werden über RFI abgewickelt. Dazu werden an bestimmten Meßpunkten die Zugverspätungen automatisch registriert. Zusätzlich werden manuell die Ursachen bzw. Verantwortlichkeiten für evtl. entstandene Verspätungen gemäß UIC-Merkblatt 450-2 [UIC09] erfaßt. Wenn kein eindeutiger Verantwortlicher bestimmt werden kann, ist per Definition der Infrastrukturbetreiber RFI verantwortlich. Für die Bonus-Malus-Zahlungen ist die Verspätung am Zielort des Zuges maßgeblich. Dazu werden abhängig von Verkehrsart und dem Zeitpunkt der Trassenbestellung Toleranzgrenzen festgelegt, unterhalb derer keine Pönale fällig werden. Der Erstattungsbetrag pro Verspätungsminute beträgt 2 €. Das letztendliche Pönal für einen verspäteten Zuglauf ist auf maximal 20 % des Trassenpreises begrenzt. Es erfolgt eine anteilige Umlegung des Erstattungsbetrages auf die einzelnen Verursacher auf dem gesamten Laufweg (vgl. Meßpunkte). Dabei werden Verspätungsminuten, die durch andere Züge desselben Verkehrsunternehmens

verursacht wurden, nicht berücksichtigt. Die Anteile, die vom verspäteten Zug selbst verursacht wurden, werden RFI gutgeschrieben. Die resultierenden Bonus-Malus-Zahlungen aller Zugläufe werden untereinander verrechnet. Einmal im Monat erfolgt ein konkreter finanzieller Ausgleich der gegenseitigen Bonus-Malus-Ansprüche.

Die Weiterentwicklung des RFI-Systems zu einem europäischen Bonus-Malus-System für den grenzüberschreitenden Verkehr ist derzeit in Diskussion. Erste Pilotanwendungen laufen laut UIC und RNE seit Mai 2010.

2.3 Eisenbahnbetriebsdisposition

2.3.1 Übersicht

Um eine gute Qualität des Eisenbahnbetriebs zu erreichen, ist eine vorausschauende Betriebsdisposition erforderlich. Da im laufenden Betrieb qualitätsmindernde Störungen aufgrund externer, technischer und betrieblicher Gründe auftreten, muß durch geeignete Maßnahmen dafür gesorgt werden, daß sich die Auswirkungen dieser Störungen auf die Betriebsqualität möglichst gering halten.

Die Betriebsdisposition kann dabei prinzipiell aus der ursprünglichen Fahrplankonstruktion abgeleitet werden. Dabei erstellt ein für die Fahrplanerstellung verantwortlicher Infrastrukturbetreiber auf Basis von angemeldeten (verkehrwirtschaftlich optimalen) Trassenwünschen einen dem späteren Betrieb als Basis dienenden konfliktfreien Sollfahrplan. Die Konfliktfreiheit des Fahrplans ist dringend notwendig zur Erreichung einer zufriedenstellenden Betriebsqualität.

Die Eisenbahnbetriebswissenschaft unterscheidet im Zusammenhang mit Fahrplänen zwischen den folgenden, bereits in 2.1.5 erläuterten Konfliktarten [Bär96]:

- Belegungskonflikte,
- Anschlußkonflikte und
- Umlaufkonflikte.

Für diesen Konstruktionsprozeß stehen rechnergestützte Werkzeuge zur Verfügung, die in der Lage sind, die Fahrplantrassen mit Hilfe des bereits in 2.1.4 erwähnten Sperrzeitenmodells [Hap59] zu verwalten und daraus abgeleitete Methoden zur Konflikterkennung beinhalten. Hierbei ist insbesondere das am Verkehrswissenschaftlichen Institut der RWTH Aachen entwickelte Verfahren FAKTUS zu erwähnen, das in Form der Software RUT-K bei der Deutschen Bahn im Einsatz ist [Brü95]. Die eigentliche Lösung der Konflikte erfolgt dabei noch manuell durch den Bearbeiter.

Erste Ansätze zur automatischen Konfliktlösung im Rahmen des Fahrplankonstruktionsprozesses finden sich im auf dem FAKTUS-Modell aufbauenden Simulationsverfahren BABSI, welches auf den im weiteren Verlauf ausführlicher beschriebenen asynchronen Simulationsverfahren basiert [Grö02].

Aufgrund der bereits genannten Störungen werden bei der betrieblichen Durchführung wiederum Konflikte in den eigentlich schon konfliktfreien Fahrplan hineingetragen. Hierbei müssen die bereits im Rahmen der Fahrplankonstruktion angewendeten Methoden nochmals angewendet werden, um einen dann erneut konfliktfreien Dispositionsfahrplan erstellen zu können, der den geforderten Qualitätsanforderungen an den Eisenbahnbetrieb erfüllt [Schw97]. Die Konfliktlösung im Dispositionsfahrplan ist aber im Vergleich zum Sollfahrplan schwieriger, da zum einen deutlich weniger Zeit für die Lösungssuche zur Verfügung steht und zum anderen der Lösungsraum dahingehend kleiner ist, daß in der zeitlichen Vergangenheit liegende Fahrplanlagen nicht mehr änderbar sind (Echtzeitproblematik). Daher sollte bereits im Rahmen der Sollfahrplankonstruktion auf eine entsprechende Robustheit desselben gegenüber Störungen geachtet werden.

Die gewählten Konfliktlösungen (Maßnahmen des Disponenten) haben entscheidenden Einfluß auf die Qualität des daraus resultierenden Eisenbahnbetriebs, insbesondere durch weitere möglicherweise resultierende Folgeverspätungen und nicht eingehaltene Anschlußübergänge. Diese „Resultate“ einer Disposition (Betriebsprotokolle) können daher mit Hilfe eines Bonus-Malus-Systems monetarisiert bewertet und verglichen werden. Über die Optimierung der Bonus-Malus-Systembewertungen hinaus sollte auch der Aspekt des möglichst energieeffizienten Betriebs als sekundäres Optimierungskriterium im Auge behalten werden.

Die Disposition kann man dabei in mehrere Teilprozesse untergliedern (vgl. z. B. [Hel04]):

- *Empfang von Störungs- und Verspätungsmeldungen (Ist-Daten):* Hierbei wird auf Mittel der Telekommunikation zurückgegriffen.
- *Fahrtzeiten- und Konfliktprognose für festgelegten (planbaren) Zeitraum:* Aufgrund des Ist-Zustandes werden die Fahrtzeiten der verkehrenden Züge entsprechend vorausberechnet und darauf basierend mögliche daraus resultierende Konflikte prognostiziert. Dabei muß insbesondere auch der mögliche Abbau von Verspätungen durch Ausnutzen von Zeitreserven im Sollfahrplan beachtet werden.
- *Klassifizierung der prognostizierten Konflikte:* Diese kann grob anhand der oben genannten Konfliktarten erfolgen. Im Einzelfall kann es aber sinnvoller sein, noch eine feinere Klassifizierung vorzunehmen.
- *Auflösung der Konflikte:* Hierbei werden geeignete (und möglichst im Sinne der Qualität und des Energieverbrauchs optimale) Maßnahmen ergriffen, die ein tatsächliches Eintreffen der prognostizierten Konflikte verhindern bzw. ihre Auswirkungen so weit wie möglich abschwächen. Mögliche Maßnahmen sind hierbei verlängerte, verkürzte oder zusätzliche Halte, Anpassungen der Fahrtgeschwindigkeiten, verlegte oder zusätzliche Überholungen und Kreuzungen, Verlegungen und Umleitungen sowie Wegfall von Zügen bzw. Einsatz zusätzlicher Züge [Jac03]. Hierbei gibt es nur wenige publizierte verbindliche Regeln (wie z. B. in den DB Netz-Richtlinien 420 [DBR420] und 408 [DBR408]). Wenn überhaupt, sind „Regeln“ dabei meistens nur informell im Kopf des jeweiligen Disponenten vorhanden und schwierig in Algorithmen zu übertragen.

- *Kommunikation der gewählten Maßnahmen* (neuer Dispositionsfahrplan). Dies erfolgt in der Regel wiederum per Telekommunikation.

Seit Beginn des 20. Jahrhunderts [Leh05] war die Disposition (genauso wie die Fahrplankonstruktion) eine vollständig manuelle Arbeit. Der Disponent erhielt hierbei fernmündliche Informationen von der „Strecke“ über aktuelle Störungen und Planabweichungen und versuchte, diese mit Hilfe eines Bildfahrplans, Bleistift und Radiergummi aufgrund seiner Erfahrungen in geeignete Dispositionsfahrpläne zu verarbeiten. Im Laufe der Zeit wurden dann aber immer mehr Teilprozesse der Disposition (teil-)automatisiert (vgl. [Jac03; Bay80; Hun93]). Zugmeldungen von der Strecke werden über automatische Schnittstellen direkt aus den sicherungstechnischen Anlagen generiert, detaillierte (sperrzeitengenaue) Fahrtzeitenprognosen können mit Hilfe von Simulationen erstellt werden, ebenfalls können diese zur Bewertung und zum Vergleich verschiedener optionaler Dispositionsfahrpläne verwendet werden. Auch ist eine automatische Prognose zukünftiger Belegungskonflikte auf Basis des Sperrzeitenmodells möglich. Konflikte stellen hier nämlich die „Überschneidungen“ von Sperrzeiten unterschiedlicher Zugtrassen dar. Auch ist die Erkennung von Anschluß- und Umlaufkonflikten bei einer guten Fahrtzeitenprognose ein relativ leichtes Unterfangen. Mit zunehmender Automatisierung und einer damit verbundenen Entlastung der Disponenten von Routinetätigkeiten wurden die entsprechend von einem Disponenten zu überwachenden Bereiche immer größer, was nebenbei auch das bei vielen kleineren Bezirken mögliche Entstehen suboptimaler Lösungen vermeidet. Ferner sind stärkere Verzahnungen zwischen Disposition dem eigentlichen Betrieb (Fahrdienstleiterebene) festzustellen, um eine einfache Umsetzung der Dispositionsfahrpläne mit kurzen Kommunikationswegen zu erreichen.

So wird die Disposition im Bereich der Deutschen Bahn derzeit von sieben regionalen Betriebszentralen [Bar01] aus durchgeführt, die zunehmend Dispositions- und operative Aufgaben örtlich vereinen [Poh02]. Übergeordnet befindet sich eine Netzleitzentrale, die insbesondere für die Überwachung langlaufender hochwertiger Fernzüge im Personen- und Güterverkehr sowie deren Anschlüsse untereinander zuständig ist [Bre03].

Am Ende der Entwicklung steht ein ausgereiftes Dispositionssystem, das alle oben genannten Teilprozesse der Disposition in sich zusammenfaßt und automatisiert für einen sicheren und wirtschaftlichen Betrieb transparent ausführt und deren Ergebnisse über entsprechende Schnittstellen zu Informationssystemen innerbetrieblich und gegenüber dem Endkunden kommuniziert. Dabei erfolgt ein unmittelbarer technischen Durchgriff eines vollautomatischen Dispositionssystems auf die Steuerungseinrichtungen der Strecken und Züge. Personal würde in diesem Falle (zumindest im Regelbetrieb) lediglich noch zu Kontroll- und Überwachungszwecken verwendet. Ausnahmen bestehen ggf. bei größeren Störfallszenarien (z. B. Sperrungen größerer Infrastrukturbereiche), wo menschliche Mitwirkung am Störfallmanagement erhalten bleiben dürfte. Die entscheidende Frage wird hierbei sein, inwieweit sicherheitsrelevante Handlungen vollständig von einem automatischen System abgedeckt werden können.

Intensive Forschungstätigkeiten im Zusammenhang mit Dispositionssystemen sind im Bereich der automatischen Konfliktlösung zu erkennen, die in der Praxis derzeit überwiegend

noch manuell erfolgt. Dadurch bedingt, daß es sich hierbei um ein NP-schweres Problem handelt [Jov89; Pet86], ist eine Lösung entsprechend schwierig zu finden. Insbesondere mit steigenden Zugzahlen steigt die Komplexität dieses Problems erheblich an. Man kann die derzeitigen wissenschaftlichen Ansätze zur Bewältigung dieses Problems grundsätzlich in zwei Gruppen unterteilen [Hel04]: Optimierungsverfahren auf Basis einer vorgegebenen Zielfunktion sowie Verfahren, die mit Hilfe vorgegebener Regeln eine Lösung erreichen.

Im folgenden sollen die Grundprinzipien einiger Ansätze aus diesem beiden Gruppen kurz dargestellt werden, ohne dabei die konkreten Verfahren detailliert zu beschreiben. Für eine eingehendere Betrachtung der einzelnen Verfahren wird in diesem Zusammenhang auf die angegebene Literatur verwiesen.

2.3.2 Konfliktlösung auf Basis einer Zielfunktion

Die Ansätze zur Konfliktlösung auf Basis einer Zielfunktion kann man grob in folgende Klassen unterteilen (vgl. [Hel04]):

- Enumerationsverfahren (Ereignisbäume), welche unterteilt werden können in
  - vollständige Enumeration,
  - Ordered Search und
  - Branch & Bound;
- genetische Algorithmen, sowie
- geschlossene Optimierungsverfahren, wiederum unterteilbar in
  - lineare Optimierung und
  - nichtlineare Optimierung.

Alle Ansätze basieren entsprechend auf der Optimierung des Verkehrs auf Basis einer vorgegebenen Zielfunktion. Hierbei ergibt sich aber bereits eine erste grundsätzliche Einschränkung für den Einsatz bei der Eisenbahnbetriebsdisposition. Um die Verfahren beherrschbar und im Rahmen der vorgegebenen Zeit berechenbar zu halten, können in der Regel nur relativ „einfache“ stetige Zielfunktionen verwendet werden, die nur eine geringe Dimensionalität aufweisen (z. B. Minimierung von Verspätungsminuten, Flüssighaltung des Verkehrs oder Minimierung folgeverspäteter Züge). Es ist schwierig, berechenbare multidimensionale Zielfunktionen zu erstellen, die gleichzeitig zueinander orthogonale (oder gar widersprüchliche) Dispositionsziele verfolgen. Außerdem ist es kaum möglich, an realen verkehrswirtschaftliche Zusammenhängen orientierte und damit komplexere Bonus-Malus-Systeme in derartigen optimierbaren Zielfunktionen abzubilden, da hier von nichtstetigen Pönalen (z. B. fixer Malus ab n Verspätungsminuten) auszugehen ist. Nichtstetige Pönale führen aber automatisch auch zu nichtstetigen Zielfunktionen, die sich in der Regel nicht für anspruchsvollere Optimierungsheuristiken eignen, die teilweise sogar eine Linearität der Zielfunktion verlangen.

Das einfachste der **Enumerationsverfahren** stellt sicherlich die vollständige Enumeration dar. Hierbei werden sämtliche praktizierbare Lösungsmöglichkeiten aufgezählt und in einem Er-

eignisbaum eingetragen. Am Ende kann man die gemäß der Zielfunktion günstigste Lösung bestimmen, die auch automatisch immer dem theoretischen Optimum (bei einer beliebigen Zielfunktion) entspricht. Praktisch kann das Verfahren auch mit Hilfe von Summationsverfahren mit Matrizen implementiert werden (vgl. z. B. Algorithmus von Floyd-Warshall zur Berechnung sämtlicher kürzester Wege in einem Graphen [Flo62]). Dieser Ansatz ist im Zusammenhang der netzweiten Eisenbahnbetriebsdisposition aber lediglich in sehr einfachen und kleinen Netzen mit sehr schwachem Verkehr denkbar, da andererseits ein extrem hoher Zeit- und Speicherplatzbedarf zur Lösungsfindung besteht, der mit den Echtzeitanforderungen der Disposition nicht vereinbar ist. Wenn dies durch „vereinfachendere“ Optimierungsstrategien reduziert werden soll, besteht wiederum die Gefahr, das eigentliche mathematische Optimum aus dem Lösungsraum zu verbannen und damit natürlich auch nicht mehr zu finden.

Aus der vollständigen Enumeration leitet sich der Ordered Search-Ansatz [Dan85] ab, der eine Mischung aus Enumeration und Suchverfahren darstellt. Dabei wird die Konfliktlösung auf das Auffinden eines Weges in einem Suchbaum zurückgeführt. Jedem Knoten (Teillösung) im Suchbaum wird eine entsprechende Bewertung zugeordnet, der die Kosten eines Lösungsweges von der Wurzel über diesen Knoten zu einem Zielknoten angibt. Die Wegsuche wird dabei grundsätzlich an dem Knoten mit der besten Bewertung fortgesetzt. Dieses Verfahren ist deutlich schneller und weniger speicherbedürftig wie die vollständige Enumeration. Bei Ordered Search kann allerdings nicht berücksichtigt werden, daß eine scheinbar ungünstigere Teillösung zu einem späteren Zeitpunkt „von alleine“ wieder besser werden kann, z. B. durch Verspätungsabbau durch Ausnutzen von Reservezeiten. Eine scheinbar ungünstigere Teillösung kann somit zu einem unter Umständen deutlich besseren Gesamtergebnis führen. Das Optimum kann daher nicht sicher bestimmt werden.

Eine weiteres Enumerationsverfahren stellt die Branch & Bound-Methode dar [Kra87; Jov91; Weg05], die eine Art Kompromiß zwischen den beiden Extrempositionen vollständige Enumeration und Ordered Search einnimmt. Man versucht, im Gegensatz zu Ordered Seach nicht nur einen, sondern mehrere Suchwege im Entscheidungsbaum zu untersuchen. Anders als in der vollständigen Enumeration werden aber wiederum auch nicht alle Wege betrachtet, sondern solche, die keine gute Lösung versprechen (z. B. Bewertung eines Zwischenknotens oberhalb einer vorgegebenen Schranke) abgeschnitten und nicht weiter betrachtet [Lan60]. Bzgl. Zeit- und Speicherplatzbedarf liegt dieser Ansatz zwischen vollständiger Enumeration und Ordered Search, gleiches gilt für die Güte des Ergebnisses. Die Wahrscheinlichkeit, das globale Optimum gemäß der Zielfunktion zu finden, ist höher als bei Ordered Search, dies ist aber immer noch nicht wirklich sicher.

Ein vollständig anderer Ansatz ist die Verwendung **genetischer Algorithmen** [Weg05, Scho06; Tak06]. Diese sind dazu geeignet, bereits vorhandene, einigermaßen „gute“ Lösungen (welche z. B. durch das Branch & Bound-Verfahren ermittelt werden können) weiterhin entsprechend in Richtung des Optimums zu verbessern. Hauptsächlich geschieht das Rekombination der vorhandenen „guten“ Lösungen (Dispositionsfahrpläne) und deren Mutation (vgl. [Mit96]). So wird z. B. der Laufweg von Zug A aus Dispositionsfahrplan X und der Laufweg von Zug B aus Dispositionsfahrplan Y in den neuen Dispositionsfahrplan Z übernom-

men. Um eine Festfahren des Algorithmus in einem lokalen Optimum zu vermeiden, werden zufallsgesteuerte Veränderungen im neu erzeugten Dispositionsfahrplan vorgenommen. Die Erzeugung neuer Lösungen auf Basis der bestehenden Lösungsmenge geht solange weiter, bis ein entsprechendes Abbruchkriterium erreicht wurde. Vorteile genetischer Algorithmen liegen darin, daß es sich um sehr schnelle Optimierungsverfahren handelt, was gerade im Bereich der Eisenbahnbetriebsdisposition von großer Bedeutung ist. Nachteile bestehen aber zum einen darin, daß nicht erwiesen werden kann, ob es sich beim Endergebnis wirklich um das globale Optimum handelt. Die Qualität des Ergebnisses hängt auch sehr stark von der Güte der Eingangslösungen ab. Zum anderen sind genetische Algorithmen bei Optimierungsproblemen mit einer hohen Zahl an den Lösungsraum einschränkenden Nebenbedingungen (wie es beim mit großen gegenseitigen Abhängigkeiten behaftete Eisenbahnbetrieb der Fall ist) nur sehr bedingt einsetzbar, da eine für genetische Algorithmen wesentliche zufällige Rekombination verschiedener Lösungen nicht ohne weiteres möglich ist. Dies läßt sich nur durch Modellvereinfachungen umgehen. Dies kann aber für die Disposition bei zu weiter Eingrenzung des Lösungsraums zur Folge haben, daß die angeblich optimale endgültige Lösung wegen vorhandener – nicht modellierter – tatsächlicher betrieblicher Einschränkungen gar nicht umsetzbar ist, oder daß bei zu enger Eingrenzung des Lösungsraums das tatsächliche Optimum nicht mehr gefunden werden kann, da es sich bereits außerhalb der Lösungsmenge befindet.

Die letzte Gruppe der Konfliktlösungsverfahren auf Basis einer Zielfunktion sind die **geschlossenen mathematischen Optimierungsverfahren**, nämlich die lineare und nichtlineare Optimierung. (Für konkrete Anwendungen siehe z. B. [Szp73; Sau83; Jov89; Kra95; Mar95]) Diese sind insbesondere bei nordamerikanischen Eisenbahnen stärker verbreitet, bei denen lange eingleisige Abschnitte und eine relativ geringe Verkehrsdichte vorherrschen. Bei der linearen Optimierung wird eine entsprechende Zielfunktion vorgegeben und der mögliche Lösungsraum durch hinreichende Nebenbedingungen eingegrenzt. Bei der nichtlinearen Optimierung werden Zwangsbedingungen und Freiheitsgrade vorgegeben, wobei die Gesamtheit der zu disponierenden Zugfahrten mit Hilfe eines Systems nichtlinearer Differentialgleichungen dargestellt wird.

Für die mathematische Lösung gibt es Standardverfahren, wie z. B. den Simplexalgorithmus (vgl. [Cor01]) für die lineare Optimierung. Diese Verfahren garantieren, daß in jedem Fall auch das globale Optimum entsprechend der vorgegebenen Zielfunktion ermittelt werden kann. Die Nebenbedingungen definieren dabei implizit die technischen und organisatorischen Rahmenbedingungen des Eisenbahnbetriebs. Hier tritt aber wiederum prinzipiell dasselbe grundsätzliche Problem wie bei den genetischen Algorithmen in verstärkter Form zutage. Die Rechenzeit zur Bestimmung des Optimums steigt mit zunehmender Anzahl der Nebenbedingungen rapide an. Für die exakte Beschreibung der vielfachen gegenseitigen Abhängigkeiten im Eisenbahnbetrieb ist allerdings ein sehr großer Satz an Nebenbedingungen erforderlich. Um den Echtzeitanforderungen der Disposition gerecht zu werden, ist es daher für praktische Anwendungen (insbesondere im dicht befahrenen Netzen) nötig, hier Vereinfachungen bei der Modellierung des Betriebs vorzunehmen. Dadurch kann der Lösungsraum wiederum mit den bereits genannten Folgen zu weit oder zu eng gefaßt werden. Die Bestimmung des Optimums kann nicht mehr sicher gewährleistet werden.

Der hohe Grad an Abhängigkeiten der Betriebsprozesse und die damit verbundenen Nebenbedingungen sowie die Tatsache, daß Kombinationen guter Teillösungen nicht zwangsläufig zu guten Gesamtlösungen führen, machen ebenfalls den Einsatz von Standard-Heuristiken, wie z. B. „Simulated Annealing“ [Kir83] oder „Tabu Search“ [Glo97] für dieses Optimierungsproblem unter den vorgegebenen Echtzeitanforderungen sehr schwierig.

2.3.3 Konfliktlösung mit regelbasierten Verfahren

Die regelbasierten Konfliktlösungsverfahren kann man allgemein unterteilen in

- simulationsgesteuerte Disposition, wiederum unterteilbar in
  - synchrone Simulationen und
  - asynchrone Simulationen, sowie
- Expertensysteme.

Hierbei ist es nicht notwendig, eine feste Zielfunktion aufzustellen, auf deren Basis die Konfliktlösungen erfolgen. Zielfunktionen können allerdings zur nachträglichen Evaluierung der Resultate bzw. des Verfahrens verwendet werden. Regelbasierte Disposition ist darauf angelegt, in möglichst kurzer Zeit gute und brauchbare Ergebnisse zu liegen. Dabei orientieren sich die Verfahren möglichst nah am eigentlichen Betriebsgeschehen und übertragen diese in entsprechende Konfliktlösungsalgorithmen. Sie führen nicht zu einer komplett neuem Dispositionspraxis, sondern versuchen, die bisher bewährten Verfahren durch Rechnerunterstützung besser und vor allem schneller zu machen.

Bei **synchronen Simulationen** (vgl. [Kla94; Hür02]) wird der Betriebsprozeß ereignisorientiert im Zeitschrittverfahren nachgebildet. Die Simulationsschritte können entweder zeitorientiert (mit festem Zeitschritt) oder ereignisorientiert (von Ereignis zu Ereignis) ablaufen. Bei festen Zeitschritten werden Ereignisse, die im Betrieb gleichzeitig ablaufen, auch im selben Simulationsschritt abgebildet. Die Simulationsschritte können z. B. durch ein Petrinetzmodell [Rei90] auf Basis von Ereignissen und Bedingungen modelliert werden. Um brauchbare Ergebnisse zu erhalten, ist eine mikroskopische (weichen- und signalgenaue) Infrastrukturverwaltung erforderlich, in der Regel als gerichteter Graph, dessen Knoten die Infrastrukturelemente abbilden. Zugfahrten können durch Belegungen der Kanten im Graphen modelliert werden. Mehrere Kanten können dabei zu Belegungselementen (z. B. Blockstrecken) zusammengefaßt werden.

Problematisch an der synchronen Simulation ist, daß Konflikte erst zum Zeitpunkt ihres Auftretens erkannt werden können. Um unerwünschte (und im Rahmen der Echtzeitdisposition nicht beherrschbare) Rollback-Operationen zu vermeiden, muß eine vorzeitige Reservierung der angeforderten Infrastrukturelemente erfolgen. Dabei gilt das Prinzip, daß Züge mit hoher Priorität eine frühzeitigere Vorausbelegung als mögliche „Konkurrenten“ erhalten. Durch Vorausbelegungen wird eine frühzeitige Konflikterkennung möglich, wodurch mehr Spielraum zur Lösungsfindung verbleibt. Es verbleibt allerdings auch bei Verwendung von Vorausbelegungen das grundsätzliche Problem, daß Probleme unter Umständen zu früh oder spät erkannt werden, was zu unnötigen zusätzlichen Wartezeiten durch voreilige oder ver-

spätete Konfliktlösungen führt. Außerdem besteht die Gefahr von Phantomkonflikten (vom Simulator erkannte Konflikte, die eigentlich gar nicht eintreffen). Ferner besteht bei eingleisigen Strecken die Möglichkeit einer Verklemmung, die ggf. vorzeitig erkannt werden muß.

Wenn die genannten Schwierigkeiten gelöst werden können, liefern synchrone Simulationen korrekte und fahrbare (aber nicht unbedingt optimale) Dispositionsfahrpläne. Bei frühzeitiger Konflikterkennung ist ein großer Lösungsspielraum vorhanden. Außerdem können Anschluß- und Umlaufkonflikte sehr gut behandelt werden und auch auf sehr kurzfristige Störungen reagiert werden.

Bei Verwendung einer Zielfunktion zur Evaluierung der Ergebnisse während der Disposition sind synchrone Simulationen methodisch sehr eng mit Enumerationsverfahren verwandt. Einen Simulationsschritt kann man hier als Pfad im Suchbaum modellieren. Wenn nur eine Simulationsrechnung auf Basis vorgegebener Standardkonfliktlösungsalgorithmen erfolgt, entspricht das dem Ordered Search-Ansatz; wenn mehrere Szenarien nacheinander simuliert werden, ist das mit Branch & Bound vergleichbar.

Die **asynchrone Simulation** [Grö02; Jac03] arbeitet demgegenüber nicht nach dem Zeitschritt-Verfahren. Das Grundprinzip ist, daß die Züge nacheinander vollständig in der Reihenfolge ihrer Priorität in das entsprechende Zeit-Weg-Diagramm des Dispositionsfahrplans eingelegt werden. Dieses Verfahren leitet sich aus dem Vorgehen bei der Fahrplankonstruktion ab. Häufig wird der Kapazitätsverbrauch von Zugtrassen wiederum mit Hilfe von Sperrzeiten (vgl. 2.1.3 bzw. 2.1.4) modelliert. Beim Einfügen einer Zugtrasse in den Dispositionsfahrplan wird entsprechend darauf geachtet, daß in allen Blockabschnitten keine Überschneidung seiner Sperrzeit mit Sperrzeiten von bereits früher eingelegten Zugtrassen erfolgt. Wie bei der synchronen Simulation ist zur Erstellung von fahrbaren Dispositionsfahrplänen wiederum eine mikroskopische Infrastrukturmodellierung in Form eines gerichteten Graphen erforderlich. Bedingt durch ein Einlegen der Züge in Reihenfolge ihrer Priorität, kann sichergestellt werden, daß ein niederrangiger Zug nicht durch einen höherrangigen Zug behindert werden kann. Für die Reihenfolge des Einlegens von Zügen derselben Prioritätenstufe gibt es unterschiedliche Ansätze.

Bedingt durch das sehr strikte Einhalten vorgegebener Prioritätenstufen entstehen bei der asynchronen Simulation teilweise „unflexible“ Fahrpläne. (Es kann in Einzelfällen insgesamt günstiger sein, einen hochprioritären Zug kurz warten zu lassen.) Außerdem ist der verkehrswirtschaftliche Wert eines Zuges auf seinem Lauf nicht immer als konstant zu betrachten, sondern Schwankungen unterworfen, die nicht mit einer globalen Prioritätenklasse erfaßt werden können. Ebenfalls ist es vergleichsweise schwierig, Anschluß- und Umlaufkonflikte geeignet zu behandeln. Darüber hinaus ist die Behandlung eingleisiger Strecken unter Umständen mit Problemen verbunden (frühzeitiger Verbrauch der gesamten Streckenkapazität durch „ungünstige“ Lage der hochprioritären Züge).

Vorteile der asynchronen Simulation liegen darin, daß garantiert (belegungs-)konfliktfreie und damit fahrbare Fahrpläne erzeugt werden, die aber nicht unbedingt im mathematischen Sinne optimal sein müssen. Dieses ist in relativ kurzer Zeit ohne unerwünschte Rollback-Prozesse oder komplexe Entscheidungsprozesse möglich. Der Algorithmus ist garantiert

verklemmungsfrei, wobei allerdings unter Umständen nicht alle Zugfahrten auf eingleisigen Strecken eingelegt werden können. Außerdem können (im Gegensatz zu synchronen Simulationen) Konflikte frühzeitig exakt detektiert werden (außer bei sehr plötzlich eintretenden Störungen), so daß in geeigneter Weise reagiert werden kann.

Wie bei synchronen Verfahren können Ergebnisse der asynchronen Simulation im Rahmen der Disposition mit Hilfe einer Zielfunktion bewertet werden. Entsprechend den Enumerationsverfahren können neben der „Standardlösung“ mehrere Varianten (mit unterschiedlichen Zugprioritäten) gerechnet werden und die beste Lösung als Dispositionsfahrplan verwendet werden.

Neuerdings sind eine verstärkte Konvergenz bei der Weiterentwicklung und damit ein Aufweichen der ehemals strikten Grenze zwischen synchronen und asynchronen Verfahren zu erkennen. In synchronen Verfahren wurden teilweise asynchrone Komponenten aufgenommen, so daß auch eine automatische Konstruktion von Soll-Fahrplänen grundsätzlich in ähnlicher Weise ermöglicht wird, wie dies bereits in asynchronen Verfahren praktiziert wird [Hau00]. Ferner gibt es Tendenzen, nur die prognostizierte Belegungszeit eines Zuges in einem Streckenabschnitt für Vorausbelegungen zu berücksichtigen, um Phantomkonflikte zu vermeiden [Hau99]. Im Zentrum der Forschung bei asynchronen Simulationen steht derzeit die derzeit noch unzureichende Integration der Behandlung von Anschluß- und Umlaufkonflikten [Scha05]. Langfristig scheint es auch denkbar, daß ein hybrides Simulationsmodell entsteht, das die Vorteile aus synchronen und asynchronen Simulationsverfahren in sich vereint.

Neben den Simulationen gibt es als weiteres Verfahren zur Konfliktlösung die Verwendung von Wissensbasen in **Expertensystemen** (vgl. allg. [Jack99], im Eisenbahnkontext auch [Freu87]). Diese sind ein wichtiger Bereich des Forschungsgebietes der Künstlichen Intelligenz. Dabei handelt es sich um rechnerunterstützte Applikationen, die nichtformalisiertes Wissen über bestimmte Probleme speichern und daraus mit Hilfe logischer oder heuristischer Verfahren Schlußfolgerungen für konkrete Problemstellungen ermitteln, die auf dieser Basis ggf. mit manueller Unterstützung gelöst werden.

Es besteht aus fünf Kernkomponenten:

- einer Benutzerschnittstelle zur Problemlösung,
- einer Inferenzmaschine zur Suche potentieller Lösungen in der Wissenbasis,
- einer Erklärungskomponente zur Begründung und Kommentierung der Lösungsfindung,
- einer Programmierumgebung zur Bearbeitung der Wissensbasis, sowie
- der aus Expertenwissen abgeleiteten Fakten und Regelbasis.

Dabei wird zwischen deduktiven und induktiven Systemen unterschieden. Dabei sind die deduktiven Systeme wesentlich stärker strukturiert als die induktiven, bei denen die Regeln aus konkreten Beispielen durch Algorithmen abgeleitet werden müssen.

Anspruchsvolle Kernaufgabe beim Aufbau eines Expertensystems ist die Formalisierung des unstrukturierten Expertenwissens in eindeutigen Fakten und Regeln. Dabei wird in der Regel zunächst mit einer Teilmenge des Wissens begonnen, welche anschließend ergänzt und optimiert werden kann.

Expertensysteme bieten sich insbesondere bei Problemlösungen an, bei denen ein umfangreiches und strukturierbares Expertenwissen effizient in einem Rechner verwaltet werden kann oder wenn es für das Problem noch keinen implementierbaren Algorithmus gibt, was auch auf die rechnerunterstützte Konfliktlösung bei der Eisenbahnbetriebsdisposition zutrifft.

Der praktische Einsatz von Expertensysteme bei der Betriebsdisposition erstreckt sich derzeit entweder auf Anwendungen unter betrieblich stark vereinfachten Rahmenbedingungen (z. B. der homogene und abgeschlossene Verkehr japanischer Shinkansen-Linien [Hyo97]) oder auf spezielle, klar abgegrenzte Teilaufgaben der Betriebsdisposition (z. B. Entscheidungsunterstützung bei der Zulassung von Fahrten auf gesperrten Streckenabschnitten [Kos90] oder die Vermeidung von Deadlocks auf Strecken mit eingleisigen Abschnitten [Pac93]). Außerdem wird vorgeschlagen, zur Modellierung der Wissensbasis für die Konfliktlösung Fuzzy-Petrinetze zu verwenden [Fay99]. Diese ermöglichen es, Wahrscheinlichkeiten für die Wirksamkeit bestimmter Dispositionsregeln zu definieren und ermöglichen damit die Verwendung „unscharfer“ Informationen.

Grundsätzlich sind Expertensysteme zur Unterstützung der Echtzeitbetriebsdisposition geeignet und wurden z. B. im Betriebszentralenkonzept der Deutschen Bahn bereits einmal prototypisch eingesetzt [Pfe95]. Mit relativ geringem Aufwand und im Vergleich zu zielfunktionsbasierten Verfahren höheren Geschwindigkeiten können komplexe Konfliktlösungsentscheidungen des Disponenten unterstützt werden. Die Voraussetzung dafür ist die Formalisierbarkeit des Expertenwissens und eine begrenzte Zahl von Lösungsalternativen. Problematisch bei größeren heterogenen Bahnnetzen ist allerdings genau diese Transformation des sehr komplexen und vagen Expertenwissens, von dem die Qualität der Ergebnisse aber entscheidend abhängt. Wenn mehrere Lösungsalternativen abgeleitet werden, ist ggf. ergänzend eine Simulation und Bewertung der Resultate mit einer Zielfunktion erforderlich. Schneller ist die heuristische Auswahl einer „vernünftigen“ (aber nicht unbedingt optimalen) Lösung.

Grundsätzlich erlauben Expertensysteme auch eine dezentrale Disposition mit Hilfe verschiedener, miteinander in Kommunikation tretender Agenten für Züge und Fahrwegelemente [Fay00; Bie04]. Wegen der dabei ungelösten Verklemmungsproblematik im Zusammenhang mit eingleisigen Abschnitten und des hohen Kommunikationsaufwandes ist dieser Ansatz in größeren Netzen aber praktisch nicht umsetzbar.

2.3.4 Mögliche Kombinationen der Konfliktlösungsverfahren

Es gibt erste Versuche, Optimierungs- und regelbasierte Ansätze in einem Hybridverfahren zu kombinieren. Dabei gibt es zwei mögliche Grundansätze.

Zum einen kann das als Hauptverfahren eingesetzte regelbasierte Verfahren durch eine Optimierung unterstützt werden. Zum Beispiel könnten Vorschläge zur Anschluß- und Umlaufkonfliktlösung zur Unterstützung einer asynchronen Simulation mit einem gesonderten Optimierungsverfahren ermittelt werden (vgl. [Scha05]). Ferner kann mittels einer nicht zeitkritischen Optimierung eine geeignete Regelbasis für ein Expertensystem generiert werden, das dann für Ad-hoc-Entscheidungen bei der Disposition verwendet wird (vgl. [Kro94]).

Der zweite Ansatz besteht darin, den Lösungsraum des hauptsächlich zur Konfliktlösung verwendeten Optimierungsverfahrens durch ein regelbasiertes System vorab einzuschränken. Dies führt zu einer entsprechenden Beschleunigung der Optimierung. Dazu können Konflikte klassifiziert werden. „Bewährte“ Lösungen aus der Vergangenheit können für die einzelnen Konfliktklassen vorgespeichert werden, so daß nur das Optimum dieser bewährten Lösungsmöglichkeiten bestimmt werden muß (vgl. [EuPa05]). Wegen der umfangreichen Nebenbedingungen im Eisenbahnverkehr und der damit verbundenen eingeschränkten Vergleichbarkeit prinzipiell ähnlicher Konflikte, ist dieser Ansatz aber nur schwierig umsetzbar.

2.4 Prioritäten bei der Eisenbahnbetriebsdisposition

2.4.1 Übersicht

Insbesondere im Zusammenhang mit den im vorausgegangenen Abschnitt vorgestellten regelbasierten Dispositionsverfahren haben Prioritäten eine herausgehobene Bedeutung.

Wenn im Rahmen einer Konfliktlösung keine Zeit mehr für umfangreiche Optimierungsberechnungen besteht, ermöglichen diese auch in komplexen Betriebssituationen zügige Dispositionsentscheidungen. Der eigentliche Optimierungsprozeß findet hier nicht bei der Konfliktlösung selber, sondern zeitunkritisch bereits im Vorfeld bei der Aufstellung der Prioritäten statt. Hierbei wird wieder zwischen Reihenfolgeregelungen auf der Strecke und Übergangs- bzw. Wartezeiten für Anschlüsse bzw. Umlaufverknüpfungen unterschieden. Ziel bei der Aufstellung von Prioritäten ist es, jeweils entsprechend den vorgegebenen Qualitätsmaßstäben im Rahmen der Disposition den planmäßigen Verkehr zügig wiederherzustellen, die Flüssighaltung des Verkehrs sicherzustellen, die Gesamtpünktlichkeit zu verbessern, sowie die Infrastruktur geeignet auszulasten [Web02] oder kurzum den Betrieb entsprechend der vorgegebenen Qualitätskriterien zu optimieren und damit ökonomische Verluste zu minimieren.

Zur konkreten Festlegung der Prioritäten gibt es nur relativ wenige Quellen. [Wun75] beschreibt relativ ausführlich die bei der Deutschen Bundesbahn zum damaligen Zeitpunkt verwendete Systematik. Die ehemalige Behördenstruktur der Bundesbahn spiegelt sich in geschwindigkeitsabhängigen eindeutigen Zuordnungen der Züge auf verschiedene Zugtypen sowie relativ klare Vorrangregelungen bzw. Wartezeitvorschriften für selbige. Analog galt dies auch für weitere europäische Staatsbahnen. Infolge der Überführung der ehemaligen Bundesbahn und weiterer ehemaliger Staatsbahnen in privatrechtliche Gesellschaften und der Trennung zwischen Infrastrukturbetrieb und dem Erbringen von Verkehrsleistungen und der damit verbundenen Forderung nach einem diskriminierungsfreien Netzzugang wer-

den diese sehr detaillierten Regelungen zunehmend reduziert und verstärkt durch allgemeinere Vorgaben oder gar Einzelfallentscheidungen der jeweils verantwortlichen Disponenten ersetzt. Den Stand der Prioritätenregelungen bei der Deutschen Bahn AG geben [Web02] und [Chr04] wieder, die wiederum auf das konzerneigene Regelwerk der Richtline 420 „Betriebszentralen“ [DBR420] verweisen.

Ergänzend wurden im Rahmen dieser Arbeit Untersuchungen aus dem Jahr 2006 über Prioritätenregeln europäischer Eisenbahnunternehmen aus dem von der Europaeischen Kommission finanzierten Projekt New Opera [NeO08] betrachtet, deren Ergebnisse entscheidend mit in das Gesamtbild einfließen sollen.

2.4.2 Prioritäten bei Anschlüssen und Umlaufverknüpfungen

Anschlüsse (und darauf abgeleitet auch Umlaufverknüpfungen) werden in der Regel mit Hilfe von Übergangszeiten und Wartezeiten priorisiert. Ziel des Infrastrukturbetreibers ist es dabei, eine hohe Zuverlässigkeit der Anschlußbeziehungen zu erreichen und Verspätungsübertragungen bei unpünktlichem Zuglauf zu reduzieren.

Im Personenverkehr setzt sich die notwendige Übergangszeit aus der für die jeweilige Verkehrsstation in Abhängigkeit zu den konkreten Halteplätzen festgelegten Mindestübergangszeit (bei DB Netz meist 5 Minuten, am selben Bahnsteig 2 bzw. 3 Minuten, je nachdem, ob sich Züge gegenüberstehen) und eines zusätzlichen Zeitpuffers (bei DB Netz 3 Minuten) zusammen. Nur Umsteigerelationen, bei denen der Zeitabstand zwischen ankommenden und abfahrenden Zug mindestens diese Übergangszeit beträgt, werden überhaupt erst disponiert und gegenüber dem Endkunden in automatischen Fahrplanauskunftssystemen kommuniziert. Damit wird das Ziel verfolgt, nicht bereits bei relativ häufig vorkommenden, geringen Verspätungen eine Anschlußdisposition vornehmen zu müssen. Allerdings steigt andererseits die Gefahr, daß ein möglicher im Umsteigebahnhof bereitstehender und zur Weiterreise geeigneter Zug (der aber in diesem Sinne kein offizieller „Anschluß“ ist) vom Reisenden wahrgenommen, aber nicht mehr erreicht werden kann, was zu geringerer Kundenzufriedenheit führt und diesem darüber hinaus auch schwer vermittelbar ist.

Unter der Wartezeit versteht man die zu vertretende Abfahrtsverspätung eines auf Anschluß wartenden Zuges, einschließlich der zu erwartenden Zugfolgeverspätung beim Befahren desselben Streckenabschnitts. Wartezeiten beginnen grundsätzlich (unabhängig von vorausgegangenen Verspätungen) mit der planmäßigen Abfahrtszeit des wartenden Zuges.

Man unterscheidet die Wartezeiten noch einmal in ausschließlich von Zuggattungen abhängigen Regelwartezeiten sowie in regional möglichen Abweichungen von denselben.

Im Bereich der Deutschen Bahn ist zunehmend ein zurückhaltender Gebrauch von Wartezeiten feststellbar. Das gilt insbesondere für den Bereich des Personennahverkehrs, wo vielerorts bereits 0 Minuten Wartezeit für alle regionalen Verkehre üblich sind. Dies hängt insbesondere mit dem Verhalten von den Verkehr bestellenden öffentlichen Aufgabenträgern zusammen, die primär an einer großen Zahl „pünktlicher Züge“ interessiert sind (und

entsprechend ihr Pönalsystem darauf ausgerichtet haben) und weniger Gewicht auf die optimale Gesamtreisekette der Endkunden legen.

Im Güterverkehr werden Anschlüsse (Wagengruppenübergänge) individuell für die Rangierbahnhöfe disponiert. Hier gibt es in der Regel ebenfalls für den Bahnhof vorgegebene Übergangszeiten zwischen Ankunft des Zubringers und Abfahrt des Aufnehmers des Wagens. Wartezeiten werden meist individuell zwischen Güterbahn und Netzbetreiber, ggf. unter Berücksichtigung der Interessen des Endkunden, abgestimmt.

2.4.3 Prioritäten der Zugfahrten

Auf Streckenabschnitten werden Festlegungen über Reihenfolgen von in einem Belegungskonflikt befindlichen Zügen aufgestellt. Diese betrieblichen Zugprioritäten dürfen nicht mit den Prioritäten bei der Trassenvergabe (Fahrplankonstruktion) verwechselt werden. So müssen Trassen eines stark vertaktetes Nahverkehrssystems in einem Ballungsraum mit hoher Priorität vergeben werden, wohingegen diese Züge bei Belegungskonflikten mit anderen Zügen nicht unbedingt vorrangig behandelt werden müssen. Im folgenden geht es ausschließlich um betriebliche Zugprioritäten.

Die bei europäischen Bahnen verwendeten Prioritätenregelungen (entsprechend derer Dispositionsphilosophien) weichen dabei, insbesondere seit Umwandlung der Staatsbahnen, sehr stark voneinander ab. Grundsätzlich sind meist nur wenige Reihenfolgeregelungen explizit als bindende Vorschrift festgelegt, weitere Regelungen sind meist nur informell und vage Vorgaben, die von Disponenten flexibel in den Konfliktfällen umgesetzt werden.

In Deutschland zählt zu den fest vorgegebenen Vorschriften der Vorrang dringlicher Hilfszüge von anderen Zügen, der Vorrang bestimmter Züge auf für sie besonders gewidmeten Streckenabschnitten (insbes. S-Bahn-Netze) sowie der Vorrang schnellerer vor langsamer Züge bei gleichwertigen Produkten. Ferner werden teure „Premiumtrassen“ im Güterverkehr mit hoher Priorität disponiert. Im Vereinigten Königreich hingegen gibt es z. B. dem entgegengesetzt keine festen Regelungen und die Prioritäten werden jeweils situationsgebunden vom Disponenten vergeben.

Bei Konflikten zwischen Zügen desselben Eisenbahnverkehrsunternehmens werden seitens des Infrastrukturbetreibers häufig Behandlungen dieser Fälle explizit mit diesem Verkehrsbetreiber vereinbart oder dieser in die Konfliktlösung explizit mit eingebunden, wobei dann von den ansonsten geltenden Regelungen abgewichen werden kann.

Ansonsten werden im europäischen Raum verschiedene Strategien (oder Kombinationen davon) angewendet, die jetzt grob im folgenden aufgezeichnet werden sollen.

1. **Priorität von Reise- über Güterzügen**
   Grundsätzlich steht hier der Gedanke im Vordergrund, daß Fahrgäste grundsätzlich wartezeitsensibler als Frachtkunden sind, was in der Vergangenheit wohl auch zutreffend war. Durch Veränderungen von Produktionsprozessen („Just in time“ – Vermeidung von unproduktiven Lagerungen) ist das heute aber nicht mehr allgemeingültig anzunehmen.

2. **Priorität von vertakteten Zubringerzügen in einen Taktknoten**
   Eine Besonderheit der Priorität von Reisezügen liegt beim Vorhandensein stark vertakteter Verkehrssysteme mit systematischen „Rundum-Anschlüssen“ in Taktknoten vor. Um eine hohe Anschlußsicherheit (und damit niedrige Gesamtreisezeit) für den Endkunden sicherzustellen, werden Zubringerzüge in diese Taktknoten mit hoher Priorität disponiert. Nachteilig an dieser Strategie ist, daß dadurch regelrechte „Nadelöhre“ für den Güterverkehr erzeugt werden.

3. **Priorität von Zügen mit hohem Trassenpreis über Zuge mit geringerem Preis**
   Diese Strategie wird in der Regel in der Weise umgesetzt, daß der Infrastrukturmanager zeitsensiblen Kunden neben Standardtrassen teure „Premiumtrassen“ anbietet, die im Konfliktfall mit entsprechend höherer Priorität disponiert werden. Hier steht der wirtschaftliche Erlös des Infrastrukturbetreibers im Vordergrund. Kunden/Züge, die einen hohen Anspruch an die Servicequalität stellen und entsprechend bereit sind, einen hohen Trassenpreis zu bezahlen, sollen im Markt gehalten werden und werden somit vorrangig behandelt. Nachteilig an dieser Strategie ist, daß die jeweiligen betrieblichen Rahmenbedingungen nicht mit einfließen. So kann z. B. die Verspätung eines „teueren“ Zuges unkritisch sein, wenn dieser in der Lage ist, diesen im weiteren Laufwegverlauf wieder aufzuholen.

4. **Priorität von schnellen Zügen vor langsamen Zügen**
   Diese Strategie basiert auf der Annahme, daß schnelle Verkehre durch Behinderungen in der Regel stärker verspätet werden als langsame Verkehre. Aber auch hier gilt, daß das im realen Bahnbetrieb insbesondere mit Blick auf die Gesamtreisezeiten nicht unbedingt der Fall sein muß (z. B. wegen Anschlußaufnahme u. a.)

5. **Priorität von Zügen auf für diesen Verkehr gewidmeten Strecken gegenüber anderen Zügen**
   Wenn der Infrastrukturbetreiber bestimmte Strecken einer bestimmten Verkehrsart widmet (z. B. S-Bahnen, Hochgeschwindigkeitsverkehr oder Güterverkehr), so genießen Züge dieser Verkehrsart Vorrang gegenüber anderen Zügen. Hierbei spielen häufig politische Rahmenbedingungen eine Rolle (z. B. Sicherstellung, daß Fördergelder nicht mißbräuchlich eingesetzt werden sollen etc.)

6. **Vorrang pünktlicher Züge gegenüber verspäteten Zügen**
   Hier wird die Erreichung einer möglichst hohen Rate pünktlicher Züge angestrebt. Ein pünktlicher Zug soll nicht durch die Verspätung eines anderen Zuges „unschuldig bestraft“ werden. Dadurch werden Diskriminierungen des Infrastrukturbetreibers weitgehend vermeiden. Problematisch daran ist natürlich, daß die mittlere Verspätung der verspäteten Züge stark ansteigt und Anschlüsse häufiger nicht eingehalten werden können.

7. **Vorrang des am meisten verspäteten Zuges**
   Bei dieser Strategie geht es darum, zunächst die extremen Verspätungen zu reduzieren und die Endkunden dieser Züge nicht durch noch weitere Dispositionsmaßnahmen zusätzlich noch mehr zu verärgern. Man erhält dadurch ein vergleichsweise hohes mittleres Verspätungsniveau, aber in Gegensatz dazu eine geringere Streuung. Problematisch sind hier die vergleichsweise hohe Rate verspäteter Züge und die damit einhergehende Unzuverlässigkeit des Betriebs.

8. **Vorrang des am wenigsten verspäteten Zuges**
   Hier findet das umgekehrte Verfahren zu 7. statt. Die am wenigsten verspäteten Züge genießen Vorrang vor allen anderen. Dabei steht die Idee im Vordergrund, möglichst viele Züge wieder in eine pünktliche Fahrlage zu überführen und damit die Rate der pünktlichen Züge zu steigern. Dies ist natürlich bei leicht verspäteten Zügen einfacher als bei starken Verspätungen. Nachteilig wirkt sich hier natürlich aus, daß die verspäteten Züge meist eine relativ hohe mittlere Verspätung aufweisen.

9. **Vorrang von Zügen, deren Verspätung sich innerhalb eines vorgegebenen Intervalls befinden, vor sonstigen Zügen**
   Dieses Vorgehen spielt insbesondere im Zusammenhang mit Bonus-Malus-Systemen eine große Rolle. Ein Zug soll möglichst nur so wenig Verspätung erhalten, daß keine (oder nur eine geringe) Pönalzahlung erfolgen muß. Wenn der zulässige Verspätungskorridor, für den keine (oder die relativ geringe) Zahlung erfolgen muß, unwiderruflich überschritten wurde, fällt die Priorität des Zuges gegenüber anderen Zügen, die sich noch innerhalb des Korridors befinden, zurück.

Teilweise können diese verschiedenen Prioritäten in den jeweiligen nationalen Dispositionsunterstützungssystemen (z. B. der deutschen Betriebszentrale) farblich dargestellt werden, so daß die Disponenten darauf aufbauend zügig Konfliktlösungen entwickeln können, die dem höherrangigen Zug möglichst keine oder nur geringe weitere Verzögerungen einbringen.

2.4.4 Besonderheiten und abschließende Bemerkungen

Abweichungen von den vorgegebenen Prioritätenregelungen sind unter besonderen betrieblichen Bedingungen (insbesondere bei Störungen) möglich. So werden bei Gleissperrungen und damit verbundenen temporären eingleisigen Abschnitten Richtungs- und Geschwindigkeitsbündelungen eingesetzt, d. h. in beiden Richtungen durchfahren mehrere Züge im „Konvoi“ mit derselben Geschwindigkeit den eingleisigen Abschnitt, um die Streckenkapazität entsprechend zu erhöhen.

Außerdem kann bei starken Verspätungen auch die Streichung eines Reisezuges (d. h. Priorität 0) durchgeführt werden, wenn in absehbarer Zeit ein gleichwertiges Ersatzangebot besteht.

Es bleibt generell festzustellen, daß europaweit der Güterverkehr (mit Ausnahme von Güterverkehrsvorrangstrecken) mit niederer Priorität betrieben wird, was in einer höheren Unzu-

verlässigkeit dieser Verkehre resultiert. Derzeit gibt es allerdings seitens der Europäischen Union Bestrebungen, dem grenzüberschreitenden Güterfernverkehr auf bestimmten Korridoren auf dem Verordnungswege eine hohe betriebliche Zugpriorität zuzuweisen, um die weitere Entwicklung des internationalen Schienengüterverkehrs in Europa zu ermöglichen [EG08].

Prioritäten können abhängig von der Strategie sowohl örtlich (z. B. beim Übergang auf gewidmete Vorrangstrecken) als auch situationsbedingt beim Eintreten bestimmter Störungen/ Verspätungen schwanken. Man kann in diesem Zusammenhang das Netz in verschiedene Prioritätenabschnitte aufteilen und Prioritäten für den Normalbetrieb sowie für bestimmte, relevante Störfälle definieren.

Schließlich verbleibt zusammenfassend festzustellen, daß alle Systeme näherungsweise mit Hilfe ihres Prioritätenansatzes versuchen, die negativen ökonomischen Auswirkungen von Betriebsabweichungen (Fahrpreisrückerstattungen und auf andere Verkehrsträger ausweichende Endkunden) auf ihre Weise zu reduzieren. Da die bisherigen Systeme aber – wie vorab geschildert – zwecks einfacher Handhabung meist nur in geringer Dimensionalität aufgestellt sind, kann dieses Ziel in der Gesamtheit bisher nur unzureichend erreicht werden. Der reale wirtschaftliche Wert einer Zugfahrt stellt sich in der Regel wesentlich komplexer und multidimensionaler dar, als dies in einem für den Disponenten einfach manuell anwendbaren Regelsatz überhaupt darstellbar ist. Dieser Mangel der derzeit verwendeten Prioritätenregeln wird auch bereits in [Wun75] und [Bär98] erwähnt.

2.5 Defizite im aktuellen Stand der Wissenschaft

Eisenbahnbetrieb, der mittels eines Bonus-Malus-Systems bzgl. seiner Qualität bewertet wird, kann entsprechend der vorangegangenen Bemerkungen mit Hilfe geschlossener Optimierungsverfahren oder mit prioritätenbasierten Verfahren disponiert werden.

Optimierungsverfahren liefern mit Sicherheit optimale Dispositionsfahrpläne, haben aber den Nachteil, daß die Ermittlung der optimalen Lösung nicht sicher innerhalb des vorgegebenen Reaktionszeitraums erfolgen kann. In derzeit verwendeten prioritätenbasierten Dispositionsverfahren erfolgt die Disposition mit auf manueller Weise vorab fest vergebenen Zug- und Anschlußprioritäten in Form von statischen Rängen und Wartezeitvorschriften. Auch dieses Verfahren hat jedoch einige Nachteile.

Bezüglich der Vergabe statischer Prioritäten entstehen zunächst einmal Probleme im Zusammenhang mit den politischen Rahmenbedingungen. Bedingt durch zunehmende Bestrebungen nach einem freien Netzzugang für Eisenbahnverkehrsunternehmen ist die Vergabe von statischen Prioritäten kritisch zu betrachten, da sich hier immer ein Diskriminierungspotential des Infrastrukturbetreibers gegenüber einzelnen Verkehrsbetrieben ergibt, deren Züge niederrangig eingestuft werden könnten. Außerdem können die fest an eine Zugfahrt geknüpften Ränge nicht eine gemäß dem zugrundegelegten Bonus-Malus-System schwankende verkehrswirtschaftliche Bedeutung einer Zugfahrt wiedergeben. Als Beispiel sei ein nur noch schwach besetzter, auslaufend verkehrender Fernverkehrszug zu erwähnen, der

aber in Teilen seines Gesamtlaufweges sehr wichtige Relationen befährt. Dadurch ergibt sich eine eher „starre“ Disposition. Schließlich müssen noch die Schwierigkeiten erwähnt werden, mit statischen Prioritäten kurzfristig eintreffende Störfälle in der Disposition geeignet zu behandeln. Wenn beispielsweise ein hochrangiger Zug eine gravierende technische Störung aufweist, so daß dieser nur noch mit geringer Geschwindigkeit fahren kann, würde dieser die hinter ihm potentiell schnelleren Züge entsprechend behindern, da eine Überholung aufgrund der Prioritätenvorgabe nicht möglich ist. Es ist daher in derartigen Fällen eine kurzfristig zu erfolgende störfallbedingte Änderung der Priorität erforderlich, was mit einer statischen Priorität natürlich nicht umzusetzen ist.

Bezüglich der manuellen Vergabe der Prioritäten entstehen in gleicher Weise Schwierigkeiten. Die manuelle Prioritätenvergabe birgt zunächst einmal die Gefahr von Willkür in sich, wenn keine klar nachvollziehbaren Kriterien für diese Vergabe definiert sind. Dadurch ergibt sich seitens des Infrastrukturbetreibers wiederum ein hohes Diskriminierungspotential. Wenn diese Kriterien aber bekannt und nachvollziehbar sind, besteht eigentlich kein Grund mehr, die Prioritäten manuell festzulegen. Statt dessen kann auch gleich eine automatische Festlegung der Prioritäten auf dieser Basis erfolgen. Ein weiteres Problem ist die längerfristige, anhand des Bonus-Malus-Systems zu erfolgende Evaluierung der mit den vergebenen Prioritäten erreichten Dispositionsergebnisse. Sollten dabei unbefriedigende Ergebnisse in Form von Pönalen für den Infrastrukturbetreiber zutage kommen, muß von diesem eine mühsame manuelle Anpassung der Dispositionsprioritäten erfolgen. Die oben genannte Argumentation für das Eintreffen kurzfristiger Störfälle betrifft auch die manuelle Vergabe. Auf derartige Störfälle muß im laufenden Betrieb sehr zügig reagiert werden, was mit einer manuellen Vergabe natürlich deutlich schlechter als mit einer automatischen Vergabe zu realisieren ist.

Um diese derzeit vorhandenen Defizite im Rahmen einer prioritätenbasierten Disposition ausgleichen zu können, wird im weiteren Verlauf der Arbeit die Entwicklung eines automatischen Prioritätenvergabe- und -optimierungssystems vorgeschlagen.

3 Verwendete Modellierungs- und Spezifikationstechniken

Ein wichtiger Aspekt bei der Entwicklung und Beschreibung eines Softwaresystems ist die konzeptuelle Modellierung des jeweiligen Anwendungsgebietes. In diesem Kapitel soll zunächst dieser in der Informatik übliche Ansatz motiviert und allgemein beschrieben werden. Anschließend erfolgen eine Darstellung der wesentlichen im Rahmen dieser Arbeit verwendeten Modellierungstechniken sowie deren konkrete Umsetzung in kommerziellen relationalen Datenbanksystemen.

3.1 Grundsätzliches zur Modellierung

In den Ingenieurwissenschaften, im Handwerk und in der Kunst ist es üblich, vor der Fertigstellung eines Werks zunächst einmal ein Modell zu erstellen, um verschiedene Entwürfe auszutesten, zu bewerten und zu vergleichen. Dies dient der Förderung des Gesamtverständnisses von Eigenschaften und Funktionalität des Werks. Da es auf nicht so relevante Details verzichtet (Abstraktion), ist die Erstellung eines Modells deutlich kostengünstiger als das eigentliche Werk.

Auch bei in Form von Software realisierten Systemen empfiehlt es sich, zunächst ein Modell zu entwickeln. Insbesondere bei sehr komplexen Systemen mit wechselseitigen Abhängigkeiten ist es häufig gar nicht mehr möglich, ohne entsprechende Vorarbeiten mit einer Implementierung zu beginnen, um ein funktionstüchtiges und effizientes Ergebnis zu erreichen. Zunächst einmal müssen unterschiedliche Sichten auf Subsysteme betrachtet (modelliert) werden, die dann schrittweise zu einem Gesamtsystem zusammengefaßt werden können.

Die Modellierung erfolgt zu verschiedenen Zwecken (vgl. [Rum93]). Ein wichtiger Aspekt ist natürlich, daß die umfassende Testung des Gesamtsystems vor dem eigentlichen Bau erfolgt. Die Entwicklung komplexer Systeme ist in aller Regel mit ziemlich hohen Kosten verbunden. Systemfehler, die erst in einer relativ späten Entwicklungsphase erkannt werden, resultieren ggf. in erheblichen Mehrkosten, da umfangreiche, als fehlerhaft erkannte Arbeitsschritte möglicherweise wiederholt werden müssen. Daher empfiehlt es sich, mögliche konzeptionelle Fehler bereits frühzeitig in einem deutlich kostengünstiger zu erstellenden Modell zu erkennen und zu beseitigen.

Ein weiterer Modellierungszweck besteht in der Verbesserung der Kundenkommunikation. Häufig bestehen Schwierigkeiten, zwischen den Systementwickler einerseits und dem Kunden und späteren Systemnutzer andererseits eine gemeinsame Sprache zur Kommunikation über die Systemanforderungen zu finden. Die Gefahr wird groß, daß zwei Experten in Ihren jeweiligen Fachgebieten einfach „aneinander vorbeireden“ und es so zu gegenseitigen unentdeckten Mißverständnissen und schwerwiegenden Fehlern im System kommt, da der Entwickler von anderen Voraussetzungen ausgegangen ist als der Anwender. Ein frühzeitig erstelltes, anschauliches Modell hilft dabei, solche Probleme zu vermeiden. Der Anwender ist in einer frühen Phase in der Lage, mit bloßem Auge zu erkennen, ob der Entwickler das Systemanwendungsgebiet auch in der Weise umgesetzt hat, wie es von ihm beabsichtigt war. Hier wird aber bereits deutlich, wie wichtig die Wahl einer geeigneten Modellierungs-

technik ist. Sie muß einerseits anschaulich und verständlich sein, um vom Anwender ebenfalls als Diskussionsgrundlage akzeptiert zu werden, andererseits muß sie aber auch in der Lage sein, die wesentlichen und „kritischen" Bestandteile des Systems sehr detailliert und zweifelsfrei zu beschreiben. Die Auswahl der geeigneten Modellierungstechnik ist daher bereits ein wichtiger Bestandteil der Modellierung selber und kann mit entscheidend für den Erfolg des Modells sein.

Ein weiterer Aspekt für eine Modellerstellung besteht darin, eine frühzeitige Visualisierung des Systems zu ermöglichen. Die Durchführbarkeit und Plausibilität eines Systems kann mit Hilfe eines Modells besser evaluiert werden. Außerdem können frühzeitig Verbesserungsmöglichkeiten erkannt und bei der Planung berücksichtigt werden.

Darüber hinaus zielt das Modell (wie bereits oben erwähnt) darauf ab, die Systemkomplexität deutlich zu verringern und die vielen Aspekte des Systems überhaupt erst verständlich zu machen. Das Modell soll dazu dienen, wirklich in einem bestimmten Zusammenhang wichtige Merkmale des Systems zu abstrahieren und sonstige Aspekte entsprechend auszublenden. In diesem Zusammenhang können viele orthogonale Sichten auf das System mit unterschiedlichen Schwerpunkten erstellt werden. Alle diese Sichten sind für sich genommen unvollständig, dienen jedoch dem Ziel, bestimmte Aspekte verständlich zu machen, um entsprechend auf Fehlentwicklungen reagieren zu können. Implementierungsdetails werden in einem geeigneten Modell z. B. ausgeblendet, da hier der Fokus von der zentralen Problemlösung weggelenkt wird.

Außerdem kann die konzeptuelle Modellierung im Nachgang einer Entwicklung auch als geeignete Dokumentation eines Systems für den Anwender dienen.

Wesentliche Grundlagen für die Softwaremodellierung wurden durch den Mathematiker TURING im Jahre 1936 in [Tur36] gelegt. Die von ihm entwickelte Turing-Maschine, bestehend aus einem unendlichen Band, einem Schreib-Lese-Kopf und den drei Operationen Lesen, Schreiben und Schreib-Lese-Kopf bewegen ist u. a. gemäß [Gan80] in der Lage, sämtliche Berechnungen eines Computers zu modellieren (Church'sche Hypothese). Darauf aufbauend wurden in den folgenden Jahrzehnten weitere Automatenmodelle entwickelt (vgl. [Hop94]).

Zeitgleich mit der starken Verbreitung objektorientierter Programmiersprachen in der Mitte der 1990er Jahre (insbesondere C++ und Java) ist auch eine stärkere Verbreitung der dazu kompatiblen objektorientierten Systemmodellierungen festzustellen. [Rum93] unterscheidet in diesem Zusammenhang zwischen Objektmodellen und dynamischen Modellen.

Im **Objektmodell** werden statische, strukturelle und datenbezogene Aspekte behandelt. Hauptaufgabe ist die Abgrenzung verschiedener „Objekte" im System mit eigener Identität, Beziehungen zu anderen Objekten, verschiedenen Attributen und Operationen. Alle weiteren Modelle bauen auf dem Objektmodell als gemeinsamer Basis auf. Das Objektmodell versucht, die Aspekte der „realen Welt" abzubilden, die für die Systemanwendung bedeutend sind. In ihren Eigenschaften und Verhalten ähnliche Objekte (gleiche Attribute, Operationen, Beziehungen und Semantik) werden dabei in Objektklassen zusammengefaßt. Diese

Klassen werden in Diagrammen dargestellt und in Beziehungen zur eigenen oder anderen Klassen im Modell gestellt. Die Terminologie im Objektmodell soll den üblichen Bezeichnungen des Anwendungsgebietes entsprechen.

Im **dynamischen Modell** werden zeitliche, verhaltensmäßige und steuerungsbezogene Aspekte des Systems betrachtet. Wichtige Komponenten des dynamischen Modells sind somit Veränderungen im Objektmodell markierende Ereignisse, Ereignisfolgen, den Kontext von Ereignissen definierende Zustände und die Organisation derselben. Es definiert Operationsfolgen bei der Steuerung des Systems in ihrem zeitlichen Zusammenhang. Dynamische Modelle werden oft mit Hilfe von Zustandsdiagrammen dargestellt. Diese zeigen verschiedene Zustände von Teilaspekten des Systems sowie die Ereignisfolgen, die Übergängen zwischen diesen Zuständen entsprechen.

Diese orthogonalen Sichten des Systems stehen in strenger Abhängigkeit zueinander. Eine Operation im Objektmodell entspricht einer Aktion im dynamischen Modell. Je nach Ziel des Modells werden entsprechende Details dargestellt oder weggelassen. In einem guten Entwurf sollten jedoch alle Modelle aus sich selbst verständlich sein und inhaltliche Kopplungen auf ein Minimum reduziert werden.

In der Informatik stehen verschiedene Techniken für Objekt- und dynamische Modellierung zur Verfügung. Für Objektmodelle sei hier stellvertretend das aus der Welt der relationalen Datenbanken stammende Entity-Relationship-Modell (ER-Modell) von CHEN [Che76] erwähnt. Im Bereich der dynamischen Modellierung ist das Konzept endlicher Automaten [Gil62, Hop94] seit Jahrzehnten bewährt.

Seit den 1990er Jahren ist der Versuch einer stärkeren Standardisierung der vorhandenen Modellierungstechniken zu einer einheitlichen Modellierungssprache festzustellen. Ursache dafür ist vor allem, daß drei Entwickler ähnlicher Modellierungsmethoden, nämlich RUMBAUGH mit OMT [Rum93], BOOCH mit Booch [Boo91] und JACOBSON mit OOSE [Jac92], seit 1995 gemeinsam bei der Firma Rational Software beschäftigt waren und dort ihre Techniken zur Unified Modeling Language (UML) [Boo99] vereinigten. UML wurde 1997 von der Object Management Group (OMG) als Standard akzeptiert. Aufgrund von Anregungen von Benutzern wurde die Sprache komplett überarbeitet und ab 2004 als UML2 neu spezifiziert [UML04].

Die UML2 enthält sechs Typen von Strukturdiagrammen (Klassendiagramm, Komponentendiagramm, Kompositionsstrukturdiagramm, Verteilungsdiagramm, Objektdiagramm und Paketdiagramm) sowie sieben Typen von Verhaltensdiagrammen (Anwendungsfalldiagramm, Aktivitätsdiagramm, Sequenzdiagramm, Kommunikationsdiagramm, Interaktionsübersichtsdiagramm, Zeitverlaufsdiagramm und Zustandsdiagramm).

Problematisch an UML2 ist allerdings, daß die Modellierungssprache mittlerweile sehr überladen ist und viele Redundanzen in den zahlreichen Diagrammtypen enthält. Dementsprechend ist ein hoher Aufwand zum Lernen und Verständnis der Sprache erforderlich, was die eingangs beschriebene Motivation zur konzeptionellen Modellierung – auf einfache Weise Verständnis für ein System zu vermitteln – allerdings droht, ad absurdum zu führen. Da man

versucht hat, sämtliche Entwicklungsumgebungen in der UML2 zusammenzufassen, ist die spätere Realisierung der Modelle auf bestimmten Plattformen mit größerem Aufwand verbunden. Insbesondere gibt es größere Diskrepanzen zu nicht-objektorientierten Entwicklungsplattformen wie relationale Datenbanken, da Relationen als zentrales Kernelement des relationalen Datenmodells z. B. nur implizit in UML2 darstellbar sind.

Aus diesen Gründen folgt diese Arbeit im folgenden nicht strikt dem UML2-Standard. Es wird versucht, entsprechend dem ursprünglichen Ziel der konzeptuellen Modellierung mit einem möglichst kleinen, für den Leser leicht verständlichen Satz von in der Informatik verwendeten und bewährten Modellierungstechniken unabhängig von deren Aufnahme in den UML2-Standard die wesentlichen Aspekte des im Rahmen dieser Arbeit entwickelten Systems nahezubringen.

Im folgenden Abschnitt sollen die in der Arbeit verwendeten Bausteine der objektorientierten Modellierung im einzelnen allgemein vorgestellt werden. Darauf aufbauend erfolgt dann ein spezieller Fokus auf Modellierungstechniken für Regeln, die als zentrale Komponente im Rahmen des in dieser Arbeit entworfenen Prioritätenvergabe- und -optimierungssystems entwickelt werden sollen.

3.2 Objektorientierte Modellierungstechniken

Allgemeines Ziel der objektorientierten Modellierung eines Systems ist eine im Rahmen der jeweiligen Anwendung hinreichend exakte Beschreibung der „Objekte“ im Kontext des Anwendungsgebietes (vgl. [Rum99]). Ein derartiges Objekt wird sowohl durch bestimmte Eigenschaften als auch ein spezifisches Verhalten genau beschrieben. Ein Objekt in diesem Zusammenhang muß kein Gegenstand „zum Anfassen“ sein, es kann sich genausogut auch um ein Konzept handeln. Vier wesentliche Begriffe sind mit dem Objektkonzept stark verbunden: Identität, Klassifikation, Polymorphismus und Vererbung.

Mit Identität wird ausgedrückt, daß Eigenschaften und Verhalten diskreten, unterscheidbaren Objekten zugeordnet werden (z. B. Zug „IC 4711“). Jedes Objekt hat seine eigene Identität und ist damit von anderen Objekten klar unterscheidbar. Unter Klassifikation wird die Gruppierung von Objekten mit gleichen Eigenschaften (Attribute) und gleichem Verhalten (Operationen) verstanden. Eine mögliche Klasse ist z. B. „Zug“ mit einem Attribut „Zugnummer“ und einer Operation „Fährt (Von, Nach)“. Die Klasse ist eine Abstraktion, die für die Anwendung wesentliche Eigenschaften beschreibt. Sie beschreibt eine (theoretisch unendliche) Menge von Objekten. Objekte sind Instanzen einer Klasse mit verschiedenen Attributwerten, wobei die Attributnamen und Operationen für alle Objekte der Klasse gelten. Mit Polymorphismus wird das unter Umständen unterschiedliche Verhalten der gleichen Operation in unterschiedlichen Klassen bezeichnet. Es kann unterschiedliche Methoden zur Implementierung der Operation geben. Mit Vererbung wird die gemeinsame Verwendung von Attributen und Operationen verschiedener Klassen auf Basis einer hierarchischen Beziehung beschrieben. Eine Oberklasse kann in verschiedene Unterklassen verfeinert werden. Die Eigenschaften der Oberklassen werden auf die Unterklassen transformiert und müssen dort nicht erneut definiert werden. Jede Instanz einer Unterklasse ist gleichzeitig auch eine Instanz der zugeord-

neten Oberklasse. Eine Bahnanlage kann z. B. eine Strecke oder ein Bahnhof sein. So können Redundanzen im Modell vermieden werden.

Wie bereits einführend bemerkt, unterscheidet man bei der objektorientierten Modellierung zwischen dem Objektmodell und dem dynamischen Modell.

3.2.1 Objektmodell

Ein Objektmodell enthält die Objekte bzw. Klassen im System bzw. einer Anwendungsumgebung, die Beziehungen zwischen diesen Objekten bzw. Klassen, sowie die Attribute bzw. Operationen der Objektklassen. Die Abbildung der realen Welt im Objektmodell muß dabei so detailliert wie nötig sein, um die Systemanforderungen genau wiedergeben zu können. Gleichzeitig soll es aber auch so allgemein wie möglich sein, um die Übersichtlichkeit und Lesbarkeit für den Anwender sicherzustellen. Mit anderen Worten, der Fokus muß auf die für das System wesentlichen Aspekte gelegt werden.

Wie bereits in den einführenden Bemerkungen erwähnt, existieren je nach Modellierungsschwerpunkt mehrere Techniken zur Darstellung eines Objektmodells. Beim im weiteren Verlauf dieser Arbeit entwickelten System spielt die Modellierung von Beziehungen zwischen Objekten beschreibenden Regeln (siehe 3.3) eine wichtige Rolle. Um dies später geeignet im Modell ausdrücken zu können, soll hier eine Modellierungstechnik zum Einsatz kommen, die eine stärkere Heraushebung von Beziehungen im Modell ermöglicht. Daher wird hier nicht der in diesem Punkt schwächere UML2-Standard verwendet, sondern das ältere, aber bewährte Entity-Relationship-Modell (ER-Modell) von CHEN [Che76]. Dieses Modell hat außerdem den Vorteil, daß es auch für in der konzeptuellen Modellierung unerfahrenen Leser sehr leicht und intuitiv verständlich ist.

Im Rahmen dieser Darstellung können nur die im weiteren Verlauf der Arbeit unmittelbar verwendeten Konzepte des ER-Modells beschrieben werden. Für eine detaillierte Einführung in die Modellierungstechniken muß auf die entsprechende Fachliteratur verwiesen werden (z. B. [Kem97]).

In einem ER-Diagramm werden die einzelnen Objektklassen als Entitytypen (Entitätstyp) und Beziehungen zwischen verschiedenen Objektklassen als Relationshiptypen beschrieben. Die einzelnen, konkreten Objekte werden als Entity und die Beziehungen zwischen diesen Objekten analog als Relationship bezeichnet.

Die Entitytypen werden als Rechtecke dargestellt, die den Namen der Objektklasse enthalten (Abb. 3-1). Sie können durch als Kreise dargestellte Attribute genauer beschrieben werden. Mit ausgefüllten Kreisen werden Schlüsselattribute zur eindeutigen Objektidentifizierung bezeichnet. Unter einem Schlüssel versteht man in diesem Zusammenhang die minimale Menge von Attributen, deren Werte das entsprechende Objekt innerhalb der Objektklasse eindeutig identifiziert. Falls es mehrere Schlüsselkandidaten gibt, wird ein geeigneter Primärschlüssel gewählt. Das klassische ER-Modell ermöglicht keine Angaben von Operationen der Objektklassen. Wo dies im Rahmen dieser Arbeit erforderlich ist, erfolgt die Angabe der Operation in Kursivschrift unterhalb des Entitytyp-Rechtecks.

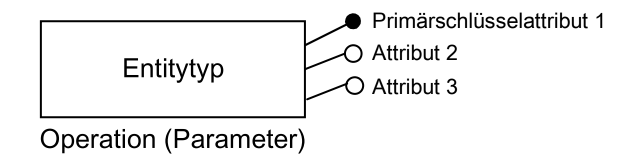


Abb. 3-1: Entitytypendarstellung im ER-Modell (Muster)

Die Relationshiptypen zwischen den Entitytypen werden als beschriftete Rauten repräsentiert (Abb. 3-2). Diesen können Attribute zugewiesen werden. Die Rauten werden mit den beteiligten Objektklassen über ungerichtete Kanten verbunden. Es besteht die Möglichkeit, durch Beschriftung der Kanten diesen Verknüpfungen Rollen zuzuordnen, was insbesondere zur besseren Klarstellung von Relationshiptypen zwischen demselben Entitytyp sinnvoll ist (z. B. ein „Zubringer“ und ein „Aufnehmer“ bei einem Relationshiptyp „Anschluß“ des Entitytyps „Zug“ mit sich selbst). Auch die Funktionalität der Beziehungen kann an den Kanten der Relationshiptypen im ER-Diagramm durch Angabe von „1“ oder „N“ erfaßt werden. Hierbei sind folgende Typen zu unterscheiden:

- 1:1-Beziehungen, wo jedem Entity e_1 des Entitytyps E_1 höchstens ein Entity e_2 des Entitytyps E_2 zugeordnet wird und umgekehrt (z. B. der Relationshiptyp „Ist Zugchef“ zwischen „1“ Bahnmitarbeiter und „1“ Zug)
- 1:N-Beziehungen, wo jedem Entity e_1 des Entitytyps E_1 beliebig viele Entities e_2 des Entitytyps E_2 zugeordnet werden aber jedem Entity e_2 des Entitytyps E_2 höchstens mit einem Entity e_1 des Entitytyps E_1 in Beziehung steht (z. B. der Relationshiptyp „Wird befördert“ zwischen „1“ Zug und „N“ Personen)
- N:1-Beziehungen (analog zu 1:N)
- N:M-Beziehungen ohne Restriktionen (z. B. der Relationshiptyp „Hält am Bahnhof“ zwischen „N“ Zügen und „M“ Bahnhöfen)

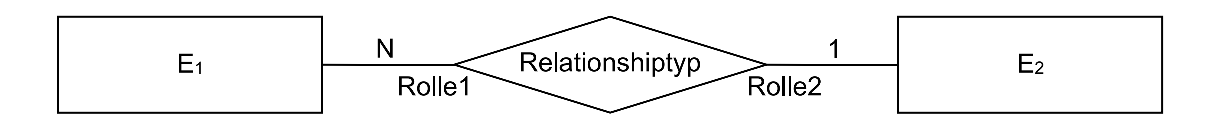


Abb. 3-2: Relationshiptypdarstellung im ER-Modell (Muster)

Diese Funktionalitäten können analog auch auf n-stellige Relationshiptypen erweitert werden. Die am Entitytyp E_k ($1 \leq k \leq n$) angegebene Funktionalität entspricht dann der Funktionalität der partiellen Funktion $R: E_1 \times ... \times E_{k-1} \times E_{k+1} \times ... \times E_n \rightarrow E_k$.

Die Generalisierung bzw. Vererbung wird durch einen besonderen Relationshiptyp im ER-Schema ausgedrückt (Abb. 3-3). Anstelle der üblichen Raute wird ein Sechseck verwendet in dem die Worte „is-a“ notiert werden. Zusätzlich werden die Kanten in diesem Fall als Pfeile in Richtung der zugehörigen Oberklasse dargestellt.

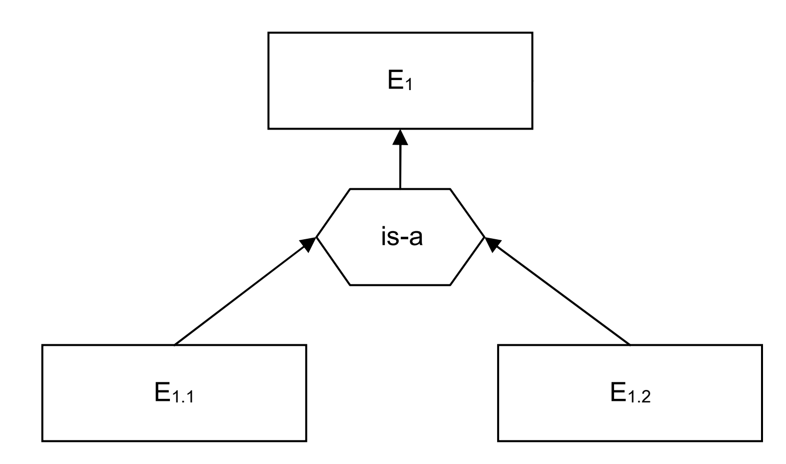


Abb. 3-3: Generalisierung/Vererbung im ER-Modell (Muster)

Ein weiterer besonderer Relationshiptyp ist die Aggregation (Teil-von-Relation/Abb. 3-4). Diese wird verwendet, um auszudrücken, daß der untergeordnete Entitytyp zwar im Gegensatz zur Vererbung eigenständige Objekte repräsentiert, dieser aber eine Komponente des obergeordneten Entitytyps darstellt (z. B. Schrauben einer Maschine). Aggregationen werden dargestellt, indem die Worte „is-part-of“ in die Raute des entsprechenden Relationshiptyps notiert werden.

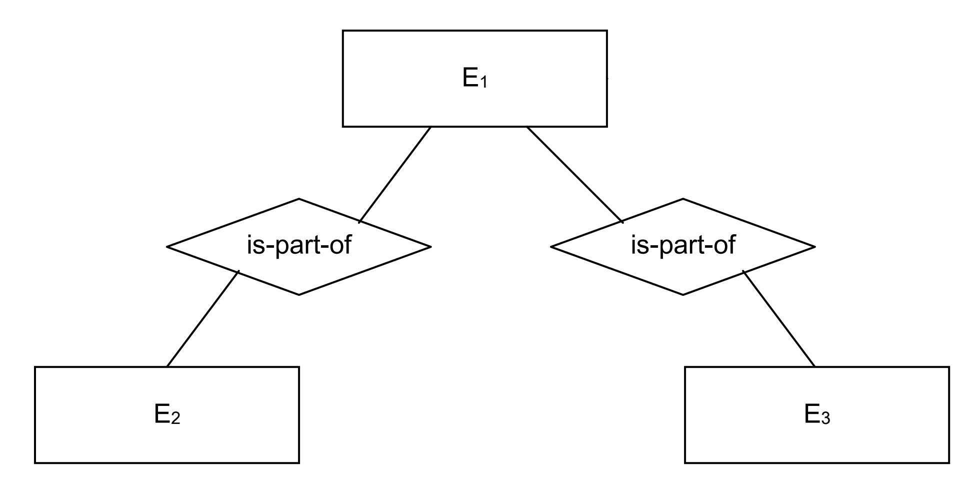


Abb. 3-4: Aggregation im ER-Diagramm (Muster)

Das Objektmodell eines zu entwickelnden Systems in seiner Anwendungsumgebung läßt sich in der Regel im erforderlichen Detaillierungsgrad nicht überschaubar, und damit für den eigentlichen Zweck der Modellierungsübung sinnvoll, in einem Diagramm darstellen. Daher ist es empfehlenswert, diese in verschiedene sogenannte Sichten aufzuteilen. Für die jeweilige Sicht wird ein spezieller Entwurf durchgeführt, der sich auf den notwendigen Bedarf beschränkt. Aufgrund der jeweiligen Anforderungen sind die verschiedenen Sichten des Entwurfes in der Regel nicht disjunkt, sondern überschneiden sich in mehr oder weniger großem Umfang. Sie können anschließend zu einem globalen Schema zusammengefaßt werden, welches redundanz- und widerspruchsfrei ist und wo die Synonyme und Homonyme bereinigt sind.

3.2.2 Dynamisches Modell

Das dynamische Modell basiert auf den strukturellen Informationen des Objektmodells. Es werden zeitliche Veränderungen von Objekten und Beziehungen illustriert. Dabei steht die Reihenfolge von aufgrund externer Reize ausgelösten Operationen unabhängig von deren konkreter Implementierung im Vordergrund.

Der Zustand eines Objektes wird durch seine Attributwerte und Verknüpfungen bestimmt. Ereignisse sind in diesem Zusammenhang Reizausübungen eines Objektes auf ein anderes Objekt. Sie geschehen in einem bestimmten Augenblick und haben in dieser Hinsicht keine meßbare Dauer. Bei einem Ereignis bzgl. eines Objektes kann je nach Objektzustand eine Zustandsänderung oder das Senden eines weiteren Ereignisses an den ursprünglichen Sender oder ein weiteres Objekt erfolgen (Aktion). In diesem Falle stehen die Ereignisse in einem Kausalzusammenhang, ansonsten gelten sie als parallel. Objekte in einem bestimmten Zustand können das Eintreten bestimmter Ereignisse erwarten, z. B. kann der am Haltsignal X stehende (Zustand) Zug IC 234 (Objekt) die Freigabe des Streckenabschnittes 1234 (Ereignis) erwarten. Analog zu Objekten können auch Ereignisse in Ereignisklassen mit gleicher Attributmenge gruppiert werden.

Ein Übergang eines Objektzustandes in einen anderen Zustand setzt ein Ereignis voraus und wird in Reaktion darauf vorgenommen. In einem bestimmten Zustand reagiert ein Objekt jeweils eindeutig auf ein Ereignis. Ein Zustand hat eine bestimmte Dauer, nämlich den Zeitraum zwischen zwei einen Zustandsübergang auslösenden Ereignissen. Ein Zustand kann mit einer andauernden Aktivität des Objektes (z. B. einer bestimmten Operation) verbunden sein. Außerdem kann er mit einem bestimmten Attributwert verknüpft sein.

Diese Ereignisse, Zustände und Zustandsübergänge können für jede bedeutsame Klasse abstrahiert und als Zustandsdiagramm (Graph aus Zuständen und Ereignissen) repräsentiert werden. Dabei konzentriert man sich auf in diesem Zusammenhang wesentliche Attribute der Objekte. Die Gesamtheit der Zustandsdiagramme wird als dynamisches Modell bezeichnet. Die einzelnen in diesen Diagrammen dargestellten Abläufe sind dann parallel und unabhängig voneinander aktiv.

Die Darstellung von Zustandsdiagrammen in dieser Arbeit ist an den Zustandsdiagrammen des UML2-Standard [UML04] angelehnt und sollen jetzt im einzelnen vorgestellt werden. Sie basieren im wesentlichen auf endlichen Automaten [Hop94], ergänzt um einzelne Aspekte klassischer Flußdiagramme [Wes71]. Wiederum muß für eine detaillierte Darstellung der Notationen auf die oben angegebene Fachliteratur verwiesen werden.

Das Zustandsdiagramm für ein Objekt beschreibt hier eine hypothetische Maschine, die sich jederzeit in einem Zustand aus einer endlichen Menge von Zuständen befindet (Abb. 3-5). Das Diagramm besteht aus

- einem Anfangszustand, der als gefüllter Kreis dargestellt wird,
- einer endlichen Menge von Zuständen bzw. Unterzuständen, die als abgerundete beschriftete Rechtecke abgebildet werden,

- einer endlichen Menge von Transitionen zwischen den Zuständen, die als beschriftete Pfeile angezeigt werden,
- einer endlichen Menge von Ereignissen, welche die Transitionen auslösen können und daher an den Transitionspfeilen notiert werden,
- Fallunterscheidungen, die als beschriftete Raute abgebildet werden und
- einem Endzustand, der als leerer Kreis mit einem kleineren gefüllten Kreis in der Mitte veranschaulicht wird.

Die Beschriftung einer Transition enthält zum einen das auslösende Ereignis mit ggf. zugehörigen Parametern in Klammern, Bedingungen (Wächter) für die Transition in eckigen Klammern und bei Auslösung der Transition durchgeführte Aktionen hinter einem Schrägstrich ggf. mit in Klammern angegebenen Parametern.

Falls ein Zustand mit der Ausführung einer Aktivität verbunden ist, wird in der Zustandsbeschriftung hinter einen „do:“ die entsprechende Aktivität mit ggf. in Klammern vermerkten zugehörigen Parametern angegeben. In diesem Zusammenhang sind auch unbeschriftete Transitionen möglich, die als eine automatische Transition aufgefaßt werden, die unmittelbar nach Ablauf der vorhergehenden Aktivität ausgelöst werden.

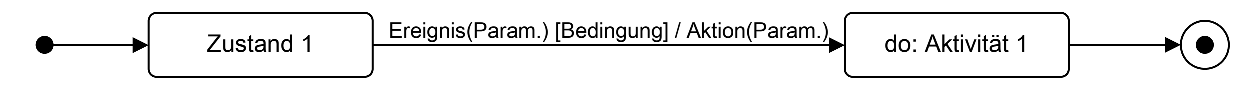


Abb. 3-5: Zustandsdiagramm mit Transition (Muster)

In Einzelfällen sind die dynamischen Prozesse eines Objektes zu komplex, um sie sinnvoll in einem einzigen Diagramm darzustellen. In diesem Falle bietet es sich an, einzelne Zustände des Diagramms in Oberzustände zusammenzufassen und Details in ein separates Unterzustandsdiagramm dieses Oberzustandes auszulagern. Diese Unterdiagramme verwenden dieselbe Syntax wie das „Haupt-Zustandsdiagramm“, wobei es komplett durch ein abgerundetes Rechteck umgrenzt wird, das mit Namen des detailliert beschriebenen Oberzustandes markiert wird. Das aus dem Anfangszustand führende Initialisierungsereignis des Unterzustandsdiagramms tritt ein, sobald das Objekt in den Oberzustand eintritt. Wenn der Endzustand des Unterzustanddiagramms erreicht wird, verläßt das Objekt gleichzeitig den Oberzustand.

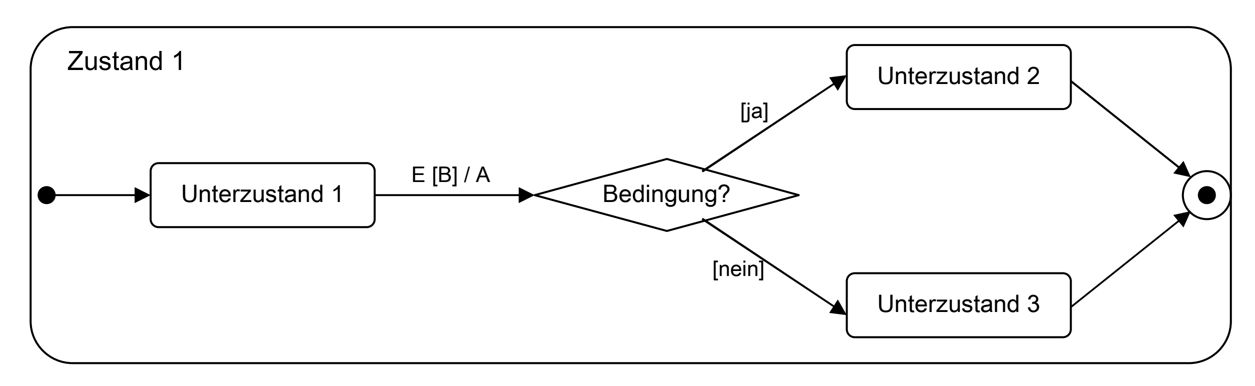


Abb. 3-6: Zustandsdiagramm mit Unterzustandsübergängen (Muster)

3.3 Regelbasierte Modellierung

Wie bereits in der Einleitung zu dieser Arbeit erwähnt, soll das hier entwickelte Prioritätenvergabe- und -optimierungssystem regelbasiert unter Verwendung von Datenbanktechnologie entworfen werden, insbesondere wegen der einfacheren Implementierbarkeit und Wartbarkeit, aber auch wegen besserer Dokumentierbarkeit Dritten gegenüber. Daher sollen in diesem Abschnitt nun ausführlicher die im einzelnen verwendeten Regelarten zur Formulierung von Zusammenhängen und Gesetzmäßigkeiten innerhalb des Systems und ihre Einbindung in die im vorigen Abschnitt eingeführten objektorientierten Modelle diskutiert werden. Hierbei stehen drei verschiedene Typen von durch „intelligente" Datenbanksysteme unterstützten Regeln im Vordergrund. Zunächst einmal die deduktiven Regeln, die dazu dienen, neue Objekte und Beziehungen aus bestehenden Objekten abzuleiten. Ferner die aktiven Regeln, die zur Definition von Systemreaktionen auf bestimmte Systemereignisse verwendet werden. Schließlich die normativen Regeln zur Sicherstellung, daß sich das System jederzeit entsprechend der vorgegebenen Semantik verhält.

3.3.1 Deduktive Regeln

Das Konzept deduktiver Regeln, also der Schlußfolgerung spezieller Aussagen aus allgemeineren Informationen, ist in der Informatik in mehreren Teilgebieten von Interesse. Zum einen ist hier der Bereich der Programmiersprachen zu erwähnen, wo Deduktion insbesondere in funktionalen Programmiersprachen [Pep06] (wie z. B. Haskell) und logischen Programmiersprachen (wie z. B. Prolog [Clo94]) anzutreffen ist. Ein weiteres Einsatzgebiet findet sich im Forschungsgebiet Künstliche Intelligenz (KI), hier insbesondere im Zusammenhang mit wissensbasierten Systemen / Expertensystemen / automatischem Schließen [Jack99]. Der letzte Bereich findet sich schließlich in der Datenbankforschung, insbesondere im Zusammenhang mit Sichten und Datenmanipulationssprachen [Ull88], [Gal84]. Obwohl es sich hier um drei ziemlich unterschiedliche Teilgebiete der Informatik handelt, sind doch im Zusammenhang mit dem Kontext „Deduktion" viele Parallelen und Überschneidungen zwischen den Ansätzen erkennbar. So läßt sich die „Logik" einer Deduktiven Datenbank in der Regel auch in einem Expertensystem oder einem Prolog-Programm ausdrücken. Da im Rahmen dieser Arbeit, wie bereits erwähnt, ein datenbankbasierter Ansatz für das zu entwickelnde System verfolgt werden soll, erfolgt die weitere Darstellung der Thematik entsprechend auch aus der Sicht der Datenbankforschung.

Ein objektorientiertes Softwaresystem besteht aus einer Menge von Objekten und Beziehungen bzw. Interaktionen zwischen diesen Objekten. Diese entsprechen den explizit gespeicherten Daten des Systems. Praktisch gibt es in jedem solcher Systeme eine Vielzahl von engen Abhängigkeiten und logischen Zusammenhängen innerhalb der Daten. Objekte, Beziehungen und Attribute, die nur unter der Voraussetzung der Existenz anderer Objekte bzw. Beziehungen existieren können, werden in diesem Zusammenhang als ableitbare Objekte, Beziehungen bzw. Attribute betrachtet. Beispielsweise läßt sich bei einer Beziehung Verkehrshalt zwischen einem Zug und einem Bahnhof das Attribut Haltezeit einfach aus der Differenz von Abfahrtszeit und Ankunftszeit ableiten. Bei einem strengen Taktfahrplan las-

sen sich beispielsweise auch große Teile des Fahrplans einfach aus sehr wenigen Basisinformationen (erster Zug, letzter Zug, Taktintervall) extrahieren.

Die Grundphilosophie des deduktiven Ansatzes ist daher, ausschließlich nicht ableitbare Basisdaten explizit im System zu speichern und alle sonstigen Daten mittels gespeicherter deduktiver Regeln und einer dem System zugehörigen Inferenzkomponente abzuleiten. In kommerziellen relationalen Datenbanksystemen wird dieses Konzept meist mittels Views in der Datenbanksprache SQL ausgedrückt (näheres dazu folgt in 3.4).

Die Vorteile des Ansatzes liegen darin begründet, daß nur vergleichsweise wenige Daten erfaßt und gewartet werden müssen. Das Risiko widersprüchlicher Datenbestände wird deutlich minimiert. Der Datenbestand ist weitestgehend selbstdokumentierend und deutlich übersichtlicher als bei einer vollständigen expliziten Speicherung. Die Wartung der abgeleiteten Daten erfolgt automatisch durch manuelle Änderung der entsprechenden Basisdaten.

Ein Nachteil besteht unter Umständen darin, daß der Datenzugriff bedingt durch zusätzlich erforderliche Aktivitäten der Inferenzkomponente etwas zeitaufwendiger ist als bei einem direkten Zugriff.

In diesem Zusammenhang sind zwei verschiedene Operationsformen der Inferenzkomponente bezüglich der Verwaltung der ableitbaren Daten möglich: Bei der anfragegetriebenen Inferenz sind die ableitbaren Daten nicht in materialisierter Form physisch im System gespeichert. Die Verwaltung erfolgt nur „virtuell“. Die Ableitung eines „virtuellen“ Datums erfolgt explizit nach bei jeder Anfrage. Es ist dann allerdings nur temporär als Antwort der Anfrage vorhanden. Bei der nächsten Ableitung desselben Datums muß die Ableitung vollständig wiederholt werden. Diese Methode empfiehlt sich bei „billigerer“ Inferenz und wenn auf das entsprechende Datum nur relativ selten durch das System zugegriffen wird.

Anders ist das bei der änderungsgetriebenen Inferenz. Hier sind die ableitbaren Daten permanent als physische Daten im System gespeichert. Die Inferenzkomponente wird ausschließlich dazu verwendet, auftretende Änderungen der Basisdaten entsprechend schnell und „fehlerfrei“ in die abgeleiteten Systemdaten fortzuschreiben. Die Ableitungen werden bei Änderungen nur einmal auf dem Datenbestand durchgeführt. Bei späteren Anfragen wird wieder auf die evtl. manipulierten physischen Daten zugegriffen. Dieses Verfahren empfiehlt sich insbesondere bei „teuren“ Ableitungen und häufigem Zugriff auf die entsprechenden Daten.

Es sind auch Mischformen zu diesem beiden Operationsformen möglich. Sogenannte materialisierte Sichten ermöglichen zum Beispiel, die Resultate einer Anfrage physikalisch zu speichern, um spätere Zugriffen entsprechend beschleunigen zu können. Dies macht insbesondere dann Sinn, wenn die Basisdaten über längere Zeit unverändert bleiben. Für jedes abgeleitete Objekt, Beziehung bzw. Attribut muß im Einzelfall je nach Verwendung im System entschieden werden, welche Inferenzmethode und damit Datenverwaltung sich entsprechend am günstigsten anbietet.

Die beabsichtigte Verwendung deduktiver Techniken ist im übrigen ein weiterer Grund dafür, im Rahmen der Objektmodellierung nicht auf den UML2-Standard zurückzugreifen, die diese – auch bedingt durch die nur implizite Berücksichtigung von Beziehungen zwischen Objekten – keine direkte Unterstützung für deduktive Konzepte bietet. Um abgeleitete Größen im Objektmodell des Systems deutlich kennzeichnen zu können, soll im Rahmen dieser Arbeit auf die Notation des erweiterten Entity-Relationship-Modells für die Modellierung deduktiver Informationssysteme (ERM^{ded}) [Rau95] zurückgegriffen werden. Dabei werden abgeleitete Konzepte mit gestrichelten Linien gekennzeichnet.

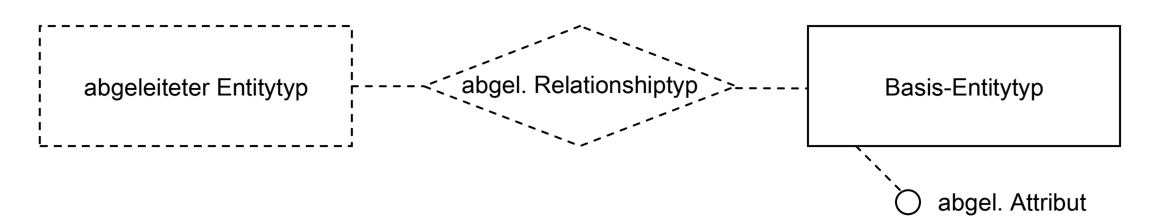


Abb. 3-8: Abgeleitete Konzepte in ERM^{ded}

Abschließend soll die Syntax und Semantik der eigentlichen deduktiven Regeln diskutiert werden. Grundsätzlich ist jede deklarative Anfragesprache zur Darstellung deduktiver Regeln geeignet. Diese Anfragesprachen basieren grundsätzlich auf elementaren Relationen-Kalkülen. Diese bestehen aus mathematischen Formeln als Sätzen, die aus elementaren Formeln und Termen zusammengesetzt werden. Die Ausdrucksfähigkeit der Kalküle ist äquivalent zur relationalen Algebra (Vereinigung, Differenz, Produkt, Projektion und Selektion elementarer Relationen). Jede Anfrage kann entsprechend auf einen derartigen Relationen-Kalkülausdruck reduziert werden. Grundsätzlich kann das Kalkül selbst daher auch als eine Anfragesprache verwendet werden. Bzgl. der verwendeten Variablen und Quantifizierungen unterscheidet man zwischen dem Bereichs- (BK) und dem Tupelkalkül (TK). Im Bereichskalkül hat jede einzelne „Komponente" der Relation eine eigene elementare Bereichsvariable. Es hat die Grundform $\{x_1, \ldots, x_n \mid f(x_1, \ldots, x_n)\}$, wobei jedes x_i eine Bereichsvariablen der dargestellten Relation und $f(x_1, \ldots, x_n)$ die Kalkülformel ist. Beim Tupelkalkül hingegen werden alle Komponenten der Relation in einer einzigen Tupelvariablen abgebildet. Es hat die Grundform $\{t \mid f(t)\}$, wobei t die Tupelvariable der dargestellten Relation und f(t) die Kalkülformel ist.

Zur Verdeutlichung der Grundprinzipien soll hier das kartesische Produkt einer n-stelligen Relation R mit einer m-stelligen Relation S in beiden Kalkülen dargestellt werden:

BK: $\{x_1 \ldots, x_n, x_{n+1}, \ldots, x_{n+m} \mid R(x_1, \ldots, x_n) \wedge S(x_{n+1}, \ldots, x_{n+m})\}$

TK: $\{t \mid \exists r \exists s\ R(r) \wedge S(s) \wedge t.A_1=r.A_1 \wedge \ldots \wedge t.A_n=r.A_n \wedge t.A_{n+1}=s.A_{n+1} \wedge \ldots \wedge t.A_m=s.A_m\}$

Für Details zu den beiden Kalkülen muß auf Fachliteratur (z. B. [Mai53]) verwiesen werden.

In der Forschung hat sich allerdings Datalog als Standardsprache für deduktive Regeln durchgesetzt. An dieser Stelle kann nur ein grober Überblick zu Datalog gegeben werden, für Details muß auf weitere Quellen (z. B. [Man08], an der sich auch diese Darstellung orientiert) verwiesen werden. Die Sprache basiert auf dem eben erwähnten Bereichskalkül, wobei

in der Grundform lediglich die Konjunktion als logische Operation der Bereichsvariablen vorgesehen ist. Daher ist „pures“ Datalog mangels einer Negation auch nur beschränkt ausdrucksfähig. In der erweiterten Form Datalog$^{\neg}$ können allerdings auch negative Literale als Bestandteile einer Konjunktion auftreten. Regeln in Datalog$^{\neg}$ sind in SQL Views überführbar (siehe dazu 3.4).

Die Syntax von Datalog orientiert sich an die Programmiersprache Prolog [Clo94], wobei Datalog im Gegensatz zu Prolog eine mengenorientierte Semantik aufweist.

Der Grundbaustein eines Datalog-Ausdrucks ist das Literal. Das Literal stellt ein „Muster“ für einzelne Fakten dar. Es besteht aus einem (kleingeschriebenen) Bezeichner und einer in Klammern angefügten zugehörigen Parameterliste. Als Parameter kommen (großgeschriebene) Variablen und (kleingeschriebene) Konstanten in Frage. Das klassische Datalog enthält keine objektorientierten Konzepte. Praktisch entsprechen die Bezeichner den Entitytypen und Relationshiptypen und die Parameter den jeweiligen Attributen eines ER-Diagramms.

Die Basisdaten eines Systems werden in Datalog als Grundliterale dargestellt. Dabei sind alle Parameter konstant, d. h. es existieren keine freien Variablen (z. B. entitytyp_1(1,’acb’) und relationshiptyp_1(1,2)). Ein Literal mit Variablen entspricht einer Anfrage nach allen möglichen Variablenbelegungen aus der Faktenmenge für das dargestellte Literal (z. B. entitytyp_1a(X,Y)).

Ableitungen von Fakten können mit Hilfe von deklarativen Datalog-Regeln definiert werden. Die Regeln haben die Form

Regelkopf $\leftarrow$ Regelrumpf.

Der Regelkopf entspricht dabei einem Muster für neu abzuleitende Fakten in Form eines Literals, meistens mit mindestens einer Variablen. Der Regelrumpf besteht aus einem Literal oder einer Kombination von Literalen.

Die Konjunktion mehrerer Literale wird als Komma zwischen denselben ausgedrückt. Für eine Regel q(X) $\leftarrow$ p(X), r(X,Y). erfolgt entsprechend während der Inferenz eine konsistente Ersetzung der Variablen in Konstanten aufgrund der Basisfaktenmenge p und r. Variablen, die nicht im Regelkopf auftreten (lokale Variablen), werden als $\exists$-quantifizierte Variablen betrachtet, mit dem $\exists$ am Anfang des Regelrumpfes. Der Bereich der Variablen im Regelkopf muß beschränkt sein, d. h. jede Variable im Regelkopf muß mindestens einmal auch im Regelrumpf auftreten, da ansonsten theoretisch unendliche Datenmengen definiert werden.

Falls bei einem Datalog$^{\neg}$–Ausdruck ein negatives Literal auftritt, wie z. B. bei der Regel q(X) $\leftarrow$ p(X,Y), $\neg$r(X,Y). dann wird dies nach dem Prinzip „negation as failure” ausgewertet, da eine direkte Ableitung aus nicht vorhandenen Fakten nicht möglich ist. Zunächst erfolgt eine Ersetzung der Variablen X und Y aufgrund der positives Literale (hier nur p(X,Y)) im Regelrumpf. Anschließend wird verifiziert, ob das Fakt im negativen Literal (r(X,Y)) mit der entsprechenden Variablenersetzung für X und Y vorhanden ist. Ist das nicht der Fall, geht man gemäß der „closed world assumption“ davon aus, daß das negative Literal falsch ist, so

daß die Ableitung (hier von q(X)) erfolgen kann. Um negative Literale in dieser Weise auswerten zu können, müssen alle Variablen eines negativen Literals vorher bereits in mindestens einem positiven Literal aufgeführt werden (stärkere Bereichsbeschränkung).

Ebenso kann eine Faktenmenge durch mehrere Regeln gleichzeitig definiert werden, z. B.

$p(X) \leftarrow q(X,Y).$
$p(X) \leftarrow s(X), t(X).$

Die Gesamtheit aller Regeln definiert in diesem Falle die abgeleitete Faktenmenge p(X) vollständig. Es handelt sich somit um eine implizite Disjunktion mehrerer Literalkonjunktionen. Der obere Regelsatz ist somit äquivalent zu

$$\forall X: p(X) \Leftrightarrow (\exists Y: q(X,Y)) \vee (s(X) \wedge t(X))$$

Im Regelrumpf müssen nicht nur Grundliterale auftreten. Auch andere abgeleitete Literale sind hier möglich. In diesem Fall gibt es komplexere Abhängigkeiten der Literale untereinander, die unter Umständen rekursiv sein können, wobei jede rekursive Regel mit mindestens einer nichtrekursiven Regel verknüpft werden muß, um einen „Ausstieg“ aus der Rekursion zu ermöglichen. Es ist nämlich nicht zulässig, einer abgeleiteten Faktenmenge q(X) einzelne Basisfakten (z. B. q(1), q(2)) explizit hinzuzufügen.

Bzgl. der Semantik eines deduktiven Regelsatzes gilt, daß die Inferenzkomponente den Regelsatz in mehreren Iterationsschritten so lange auf den Basisfakten und bereits in vorigen Iterationsschritten abgeleiteten Fakten durch Variableninstantiierungen auswertet, bis in einem weiteren Iterationsschritt keine weiteren neuen Fakten mehr abgeleitet werden können. Dieses „Resultat“ entspricht dann der Gesamtfaktenmenge und wird der kleinste Fixpunkt der jeweiligen Inferenzfunktion genant.

Für die Basisdaten p(1) und q(1,2) mit den Regeln

$r(X) \leftarrow p(X).$
$s(X) \leftarrow q(Y,X).$
$s(X) \leftarrow r(X).$

werden im ersten Iterationsschritt die Fakten r(1) und s(2) abgeleitet. Im zweiten Iterationsschritt erfolgt die Inferenz des Faktums s(1). Anschließend erfolgen keine weiteren Ableitungen mehr, d. h. der kleinste Fixpunkt der Inferenzfunktion ist somit die Faktenmenge {p(1), q(1,2), r(1), s(1), s(2)}.

Für eine Notation der deduktiven Regeln innerhalb der objektorientierten ERM^{ded}-Schemata wurde in [Rau95] anstelle von Datalog die Verwendung einer auf dem Tupelkalkül basierenden Abfragesprache ERC (Entity-Relationship Calculus) vorgeschlagen. Sie ist in der Lage, im Gegensatz zu Datalog Objekt-Attribut Zusammenhänge exakt darzustellen. Diese Sprache hat im wesentlichen die Zuordnung eines Mengenausdrucks für jedes ableitbare Konzept als Kernkomponente. Insbesondere bei mehreren Ableitungsmöglichkeiten können hier allerdings sehr komplexe Konstruktionen entstehen.

Daher wurde von MANTHEY die in [Eif06] vorgestellte Abfragesprache ER-Datalog aus dem klassischen Datalog und einzelnen „objektorientierten Konzepten“ aus ERC weiterentwickelt. In ER-Datalog werden als Parameter ausschließlich Objektvariablen verwendet, die entsprechend die Identitäten einzelner Objekte im Modell annehmen können. Um Attribute einzelner Objekte darstellen zu können, wird statt dessen die aus dem Tupelkalkül bekannte Punktnotation verwendet. Ansonsten bleibt die bereits vorgestellte Datalog-Syntax erhalten.

Damit ergibt sich die folgende Beispielsyntax für abgeleitete Entitytypen:

e3(X) ← e1(X), X.attribut_1 >10.
e4(X) ← e1(X), e2(Y), r(X,Y).

Aus dieser Notation ergibt sich, daß einstellige Prädikate grundsätzlich die Entitytypen, wohingehend mehrstellige Prädikate die Relationshiptypen repräsentieren.

Da Attributnamen im Objektmodell nicht automatisch eindeutig sind, sondern in unterschiedlichen Entitytypen bzw. Relationshiptypen gleichermaßen verwendet werden dürfen, kann zur Sicherstellung der Eindeutigkeit noch der Disambiguierungsparameter # nach dem Attributsnamen verwendet werden. In diesem Fall bedeutet X.a#e den Wert des Attributs a im Entity X, das zum Entitytyp e gehört. (X,Y).a#r ist der Wert des Attributs a der Relationship zwischen den Entities X und Y vom Relationshiptyp r.

ERM^{ded} verbunden mit ER-Datalog soll auch im weiteren im Rahmen dieser Arbeit zur Darstellung deduktiver Konzepte im Objektmodell verwendet werden.

3.3.2 Aktive Regeln

Das Konzept aktiver Regeln, also die Definition automatischer Systemreaktionen beim Eintreten bestimmter detektierter Systemereignisse, ist in der Informatik wiederum in denselben Teilgebieten von Bedeutung, wie dies auch schon bei den deduktiven Regeln der Fall war. Im Bereich der Programmiersprachen finden sich das Forschungsgebiet der ereignisorientierten Programmierung [Fai06], die – meist kombiniert mit objektorientierten Konzepten – insbesondere im Zusammenhang mit graphischen Benutzeroberflächen eine große praktische Bedeutung erlangt hat. Im KI-Bereich sind sie im Zusammenhang mit produktiven Systemen zur Aktualisierung einer Faktenbasis von Interesse [Bro85]. Schließlich werden sie im Bereich der Datenbankforschung im Zusammenhang mit Triggern untersucht [Dit00], [Man97]. Auch hier gibt es viele Parallelen und Überschneidungen zwischen den verschiedenen Sichtweisen. Wiederum erfolgt aus den bereits genannten Gründen die weitere Darstellung aus Datenbanksicht.

Wie bereits gesagt, dienen die aktiven Regeln dazu, Reaktionen eines datenbankbasierten Systems aufgrund bestimmter Ereignisse und Zustände im System zu definieren. Aktive Regeln spezifizieren in diesem Sinne die Zustandsübergänge im dynamischen Modell des Systems und können als solche in den Zustandsdiagrammen aufgenommen werden. Sie sind nicht mit deduktiven Regeln zur Definition von Datenmengen vergleichbar. In kommerziellen relationalen Datenbanksystemen werden sie in der Regel durch Trigger in der Datenbanksprache SQL ausgedrückt (näheres dazu folgt in 3.4).

Ziel ist es, mit Hilfe eines aktiven Regelsatzes das System in die Lage zu versetzen, auf bestimmte, in der jeweiligen Anwendungsumgebung „kritische“ Situationen selbständig geeignet zu reagieren. In einem Eisenbahndispositionssystem stellt beispielsweise die Verspätung eines Zuges einen solchen kritischen Zustand dar. In einem passiven System muß der Anwender selber die Situation erkennen und geeignet reagieren. In einem aktiven System soll diese Tätigkeit durch das System selbständig geschehen.

Der Grundansatz eines aktiven Systems besteht dementsprechend darin, anwendungsspezifisch definierte Situationen zu erkennen und daraufhin anwendungsspezifisch definierte Reaktionen auszulösen. In diesem Zusammenhang enthält das System eine aktive Komponente, die für die Erfüllung der entsprechenden Aufgaben verantwortlich ist.

Die dazu verwendeten aktiven Regeln sind nach dem sogenannten ECA-Paradigma (E-vent – Ereignis, C-ondition – Bedingung, A-ction – Aktion) aufgebaut (Beispielsyntax orientiert sich an [Dit00]):

```
ON Ereignis(Parameterliste)
IF Bedingung(Parameterliste)
DO Aktion(Parameterliste).
```

Damit werden Hintergrundaktivitäten definiert die durch bestimmte Abläufe des Oberflächenprozesses ausgelöst werden können, wobei letzterer suspendiert wird. Hier tritt auch der wesentliche Unterschied zu einer normalen Prozedur zutage. Diese wird explizit und direkt durch ein Anwendungsprogramm ausgelöst, wohingegen dies bei aktiven Regeln nur implizit und indirekt geschieht, wobei die Ereignisüberwachung permanent im Hintergrund erfolgt.

Es existiert ein breites Spektrum aktiver Regelsprachen, die im Rahmen dieser Arbeit nicht im einzelnen dargestellt werden können. Grundsätzlich findet sich in allen Sprachen das bereits erwähnte ECA-Paradigma wieder. Allerdings gibt es kein einheitliches Verständnis für die einzelnen Komponenten und ihre Beziehungen untereinander. Dementsprechend gibt es unterschiedliche Ansätze zur Syntax und viele verschiedene Erweiterungen zur Spezifikation der Ausführungskontrolle.

Bzgl. der Ereignisse wird zwischen primitiven Ereignissen, wie z. B. Objekt-/Datenmanipulationen, Methoden-/Prozeduraufrufen und -abschlüssen, bestimmten vordefinierten Signalen und Zeitereignissen, sowie komplexen Ereignissen, die aus der Kombination verschiedener primitiver Ereignisse entstehen können, unterschieden. Ein Ereignis bezeichnet grundsätzlich ein bestimmtes, für die Anwendung relevantes Geschehen zu einem klar definierbaren Zeitpunkt, welches seitens der aktiven Komponente des Systems eigenständig erkannt wird bzw. das diesem (über „Sensoren“) mitgeteilt werden kann.

Die Bedingungen stellen sich in erster Linie als Datenabfragen bzgl. Objektzustände dar. Aber auch Anwendungsprogramme bzw. Prozeduren mit booleschem Rückgabewert sind möglich. Die Syntax orientiert sich entsprechend an die im System verwendete Anfragesprache.

Die Aktionen stellen sich in der Regel als Objekt-/Datenmanipulationen dar. Aber auch einzelne Prozedur-/Methodenaufrufe, Folgen von Anweisungen, beliebige Programmanweisungen oder Aufrufe beliebiger Programme sind denkbar. Die Syntax leitet sich meist aus der prozeduralen Sprache des Systems ab.

In einigen aktiven Regelsprachen ist die Ereignisspezifikation parametrisiert, in anderen erfolgt sie ohne Parameter. In den parametrisierten Spezifikationen werden die Parameter dann über Variablen an Bedingung und Aktion weitergegeben. Auch für die Bedingung gibt es unterschiedliche Rollen. Sie kann der Ereignisbewertung und der Parametergewinnung für die Aktion dienen. Häufig ist die Bedingung implizit Teil der Aktion, seltener implizit Teil des Ereignisses. Im Rahmen in dieser Arbeit soll die oben erwähnte ON … IF … DO … -Syntax zur Modellierung der aktiven Regeln verwendet werden.

Auch die Semantik einer aktiven Regel ist in den vielen Regelsprachen nicht standardisiert. Hier gibt es verschiedene Sichtweisen. Implizit kann sie wie folgt zusammengefaßt werden: Wenn ein durch den „E-Teil" einer aktiven Regel spezifiziertes Ereignis auftritt, dann soll für jede Antwort des „C-Teils" (Bewertung und Parametergewinnung) der „A-Teil" ausgeführt werden.

Die Variablen in E können – müssen aber nicht – in C und A auftreten. Ferner können C und A gemeinsame Variablen erhalten (vgl. Doppelrolle der Bedingung). Wegen der Sicherheit von Aktionen müssen alle Variablen in A mindestens einmal in E oder C vorkommen. Man kann die Variablen der Bedingung in Eingabe-, Ausgabe- und lokale Variablen unterklassifizieren.

Dabei gilt: Eingabevariablen = vars(C) $\cap$ vars(E),
Ausgabevariablen = vars(C) $\cap$ (vars(A) \ vars(E)),
Lokale Variablen = vars(C) \ (vars(A) $\cup$ vars(E)).

Wegen der benötigten Bereichsbeschränktheit müssen die lokalen und Ausgabevariablen in mindestens einem positiven Literal auftreten.

Allgemein gilt, daß ein Regelverarbeitungsschritt an einem Regelausführungspunkt gestartet wird. Dieser Schritt besteht aus drei nacheinander ablaufenden Phasen. Zunächst erfolgt die Triggerphase, bei der die seit dem letzten Regelausführungspunkt aufgetretenen Ereignisse inkl. der dadurch ausgelösten Regeln (ER-Paare) zu den bereits vorher vorhandenen ER-Paaren „registriert" werden. Darauf folgt die Auswahlphase (Select-Phase), in der die in diesem Verarbeitungsschritt auszuführenden ER-Paare selektiert werden. Die letzte Phase ist die Reaktionsphase (kombinierte Evaluate- und Execute-Phase). Hier werden die selektierten ER-Paare ausgeführt, also die Regelbedingungen ausgewertet (Evaluate) und die Aktionen ausgeführt (Execute).

Innerhalb dieses algorithmischen Rahmens gibt es allerdings einige Teilschritte, die nicht eindeutig spezifiziert sind. Zunächst stellt sich die Frage nach der Regelverarbeitungsgranularität, d. h. wann ein Regelverarbeitungsschritt gestartet wird. Dies kann nach jedem Systemereignis erfolgen, aber auch erst nach einer Reihe vor Ereignissen. Im ersteren Fall

spricht man von der instanzorientierten Regelverarbeitungsgranularität, die insbesondere bei zeitkritischen Anwendungen von Bedeutung ist. Im zweiten Fall wird von der mengenorientierten Regelverarbeitungsgranularität gesprochen, bei der mehrere Ereignisse in einer Transaktion zusammengefaßt werden, an deren Ende die Regelverarbeitung startet. Damit kann sichergestellt werden, daß bestimmte Aktionen zunächst einmal „ungestört" von der aktiven Systemkomponente abgeschlossen werden können, und sich das Gesamtsystem bei der Regelverarbeitung in einem stabilen und konsistenten Zustand befindet.

Der nächste nicht eindeutig spezifizierte Teilschritt ist die Frage nach der Ausführung und den Selektionskriterien der ER-Paare. Es sind dabei verschiedene Regelausführungsmodi möglich, zum einen eine strikte Trennung von Auswertung und Ausführung, zum anderen eine Durchmischung der beiden Phasen. Bei der strikten Trennung werden zuerst alle erfüllten Bedingungen der Regeln der „registrierten" Paare ausgewertet und erst danach für alle gefundenen wahren Aussagen die Aktionen ausgeführt. Damit erfolgt eine Auswertung auf ein und demselben Systemzustand und alle ER-Paare werden gleichbehandelt.

Die Durchmischung der Phasen ist etwas komplizierter und hängt von der Struktur der Menge der ER-Paare ab. Wenn mehrere Ereignisse dieselbe Regel auslösen können, muß zunächst einmal eines dieser Ereignisse ausgewählt werden. Dies kann gemäß verschiedener Kriterien erfolgen, z. B. dem Zeitpunkt des Ereignisauftretens (FIFO bzw. LIFO) bzw. der Priorität des Ereignisses. Auch eine nichtdeterministische Auswahl ist möglich. In gleicher Weise muß eine Regel ausgewählt werden, wenn mehrere Regeln durch ein Ereignis ausgelöst werden könnten. Dies kann z. B. erfolgen aufgrund des Zeitpunkts der Regelaktivierung (FIFO bzw. LIFO), der Reihenfolge der Regeldefinitionen, anhand benutzerdefinierter Prioritäten oder mit einer nichtdeterministischen Auswahl einer Regel. Die Regelauswahl kann ggf. auch vor der Ereignisauswahl erfolgen. Dann wird für das so ausgewählte ER-Paar zunächst einmal eine Bedingungsauswertung vorgenommen und die Aktion für die entsprechenden Parameter ausgeführt. Anschließend wird für das Paar die nächste Erfüllung für die Bedingung auf dem möglicherweise manipulierten Systemzustand ausgewertet usw. Beendet wird das Verfahren für dieses Paar, sobald keine neue Antwort für die Bedingung gefunden wird. Die dahinterstehende Motivation ist die Überlegung, daß eine Regelbedingung dann wahr sein soll, wenn die Aktionsausführung startet.

Schließlich sind verschiedene Kopplungsmodi zwischen der Ereigniserkennung (E), Bedingungsauswertung (C) und Aktionen (A) möglich, wobei die Kopplung zwischen E und C sowie zwischen C und A verschieden betrachtet werden müssen. Es gibt zunächst die Möglichkeit einer direkten (immediate) Kopplung, bei der C unmittelbar nach dem Regelausführungspunkt bzw. E ummittelbar nach der Bedingungsauswertung gestartet wird. Dies ist die natürlichste Kopplungsform. Ferner gibt es die verzögerte (deferred) Kopplung, bei denen C bzw. A am Ende der aktuellen Transaktion ausgeführt werden. Motivation hierfür ist, daß dieses auf Basis eines konsistenten Systemzustandes erfolgt. Schlußendlich kann C bzw. A auch entkoppelt (decoupled) in einer gesonderten Transaktion erfolgen, was sich insbesondere bei sehr komplexen Ereignisreaktionen anbietet.

Die Semantiken für die im Rahmen dieser Arbeit vorgestellten aktiven Regeln werden im jeweiligen Zusammenhang diskutiert.

3.3.3 Normative Regeln

Das Konzept normativer Regeln dient der Definition und Prüfung von konsistenten Zuständen eines Systems und findet sich wieder in allen bereits zu den anderen Regeltypen genannten Teilgebieten der Informatik. Im Rahmen der Programmiersprachen spielen sich einerseits im Zusammenhang mit Zusicherungen (Asserts) eine Rolle [Hoa69], andererseits im Zusammenhang mit Constraintprogrammierung (meist in Kombination mit logischer Programmierung) [Hof07]. In der KI-Welt wird die Constraintprogrammierung auch als Alternative zu traditionellen Expertensystemen für wissensbasierte Systeme untersucht [Din06]. In Datenbankkontext werden sie als Integritätsbedingungen verwendet [Kem97]. Die weitere Darstellung erfolgt wiederum aus Sicht der Datenbankforschung.

Normative Regeln bzw. Integritätsbedingungen dienen dazu, zu jedem Zeitpunkt die „semantische Integrität“ bzw. logische Korrektheit eines datenbankbasierten Systems bzgl. seiner Anwendungsumgebung sicherzustellen. Das System soll im Rahmen seiner Operationen nur sinnvolle und bzgl. der Semantik legale Zustände annehmen. In kommerziellen relationalen Datenbanksystemen werden sie in der Regel durch Constraints in der Datenbanksprache SQL ausgedrückt (näheres dazu folgt in 3.4). Als Bestandteil von UML existiert die Object Constraint Language (OCL) [OCL06], die es ermöglicht, Integritätsbedingungen in UML Modellen zu spezifizieren.

Viele normative Vorgaben sind bereits inhärent im Objektmodell des Systems spezifiziert. Dazu gehören z. B. Funktionalitäten von Beziehungen, zulässige Objektklassen in Beziehungen, zulässige Attributwerte etc. Diese Vorgaben müssen bei einer Systemimplementierung natürlich geeignet berücksichtigt werden. Im Rahmen der Systemmodellierung reicht es allerdings aus, die Aufstellung normativer Regeln auf „komplexere“ Semantiken der Anwendungsumgebung zu beschränken, die nicht bereits unmittelbar aus dem Objektmodell hergeleitet werden können. Ziel ist es schließlich, das Modell übersichtlich zu halten und auf unnötige Redundanzen zu verzichten. Im Eisenbahnkontext können das u. a. technische, organisatorische, rechtliche oder unternehmenspolitische Regeln sein.

Eine technische normative Regel eines Dispositionssystems ist z. B., daß sich nicht zwei Züge gleichzeitig in derselben Blockstrecke befinden dürfen. Eine unternehmenspolitische Regel kann sein, daß ein Zug maximal fünf Minuten auf Anschlußreisende warten darf.

Man kann in diesem Zusammenhang auch zwischen „harten“ und „weichen“ normativen Regeln unterscheiden. Integritätsbedingungen im eigentlichen Sinne sind nur die harten Bedingungen, denn diese dürfen unter keinen Umständen zu irgendeinem Zeitpunkt verletzt werden. Dazu gehört beispielsweise die oben genannte Blockregel im Dispositionssystem, da das Verletzen dieser Regel im Rahmen der Disposition kategorisch ausgeschlossen ist. (Im realen Betrieb ist ein Verletzen der Regel aufgrund technischen und menschlichen Versagens natürlich wiederum möglich, so daß die Blockregel im Rahmen eines Betriebs-

überwachungssystems keinen harten Charakter haben kann.) Eine weiche normative Regel hingegen sollte jederzeit erfüllt sein, kann aber in Einzelfällen, wenn keine andere Lösung möglich ist oder jede andere Lösung mit nicht vertretbaren Kosten verbunden wäre, mit ausdrücklicher Zustimmung des Systemadministratoren verletzt werden. Dazu gehört z. B. die oben genannte 5-Minuten-Anschlußregel. Dieses unternehmenspolitische Ziel sollte zwar prinzipiell immer eingehalten werden. In einzelnen Härtefällen sollte der Disponent allerdings in der Lage sein, von dieser Vorgabe abzuweichen, ohne daß dies vom System als unzulässig zurückgewiesen wird.

Man kann ebenso zwischen der statischen und der dynamischen Integrität eines Systems unterscheiden. Während die statische Integrität die Korrektheit des aktuellen Systemzustandes darstellt, beschreibt die dynamische Integrität die Korrektheit des Systemverhaltens (insbesondere im Zusammenhang mit Zustandsänderungen). Hier kann man ebenso die Blockregel wieder der statischen Integrität zuordnen, wohingegen die 5-Minuten-Anschlußregel einen eher dynamischen Bezug hat. Letztendlich sind die Übergänge zwischen statischen und dynamischen normativen Regeln allerdings als fließend anzusehen. Eine statische Regel kann einfach in eine dynamische Regel transformiert werden und umgekehrt. Zum Beispiel läßt sich die Blockregel dynamisch so formulieren, daß ein Zug nicht in eine Blockstrecke einfahren darf, die bereits von einem anderen Zug belegt ist.

Formal sind statische normative Regeln (und damit auch die „äquivalenten“ dynamischen Regeln) nichts anderes als geschlossene Formeln des Tupelkalküls. Sie stellen eine Definition bzw. Deklaration der erlaubten und konsistenten Systemzustände dar. Damit zeigt sich allerdings auch eine starke Verwandtschaft zu den bereits beschriebenen deduktiven Regeln. In diesem Sinne können die normativen Regeln als eine spezielle Form von „deduktiven“ Testabfragen spezifiziert werden, die grundsätzlich entweder das Fakt „true“ oder „false“ ableiten (vgl. [Man08]).

Dazu wird in diesem Zusammenhang auch im Rahmen dieser Arbeit eine an Datalog angelehnte Syntax für die Formulierung der normativen Regeln verwendet. Die normative Regel beginnt mit dem Schlüsselwort *constraint*, dem sich ein Datalog-Grundliteral anschließt. In diesem Fall muß das Fakt immer Bestandteil der Basis- bzw. abgeleiteten Faktenmenge sein. Steht zusätzlich das Schlüsselwort *not* vor dem Grundliteral, darf sich das Fakt niemals im der Gesamtfaktenmenge des Systems befinden:

```
constraint q(a).
constraint not q(a).
```

Wegen der beschränkten Ausdrucksfähigkeit von Datalog müssen für praktische Anwendungen zusätzliche Prädikate, wie z. B. Vergleiche und arithmetische Grundoperationen im Rahmen der Regelmodellierung verwendet werden.

Grundsätzlich werden auf diese Weise nur die echten (harten) normativen Regeln modelliert. Es ist jedoch einfach, weiche normative Regeln in harte normative Regeln zu überführen, indem zur Deklaration der weichen normativen Regel noch ein zusätzliches negatives Literal

überprüft wird, ob eine Administratorgenehmigung zum Brechen einer normativen Regel vorliegt:

```
harte_r(X) ← weiche_r(X), ¬brechen_ok(X).
constraint harte_r(X).
```

Damit wird implizit deklariert, daß eine weiche normative Regel bei fehlender Administratorgenehmigung durch das System wie eine harte normative Regel gehandhabt werden soll.

Es gilt allerdings das Prinzip, daß normative Regeln nur beim kompletten Abschluß einer sogenannten Transaktion, d. h. beim Erreichen eines konsistenten Systemzustandes erfüllt sein müssen. Die einzelnen (möglicherweise in zufälliger Reihenfolge) erfolgenden elementaren Teilschritte einer Transaktion dürfen im inkonsistenten Zustand zu einer temporären Verletzung der Regel führen, die allerdings im Rahmen der weiteren Teilschritte bis zum Ende der Transaktion wieder korrigiert sein sollte.

Die Überprüfung der normativen Regeln kann entsprechend dieser Überlegungen vollständig durch eine bereits vorhandene Inferenzkomponente des Systems erfolgen (vgl. Ausführungen zu deduktiven Regeln).

Die eigentliche Integritätsüberwachung selbst kann sehr gut durch eine vorhandene aktive Komponente des Systems erfolgen. Grundsätzlich kann man für jede normative Regel im System eine Menge aller möglichen Ereignisse definieren, die unter Umständen zu einer Verletzung der normativen Regel führen können. Basierend auf diese bzgl. der normativen Regel „kritischen" Ereignisse können aktive Regeln definiert werden, die gemäß einer mengenorientierten Regelverarbeitungsgranularität am Ende der Transaktion die Einhaltung der normativen Regel überprüfen und bei Verletzung ein komplettes Rollback der Transaktion anordnen können:

```
ON kritisches_Ereignis(X)
IF not (constraint p(X))
DO rollback.
```

Grundsätzlich sollen Rollback-Aktionen bei automatischen Abläufen des Systems vollständig vermieden werden. Das System sollte von vorneherein so „gebaut" sein, daß auf alle „kritischen" Ereignisse nur in einer Weise reagiert wird, die nicht zu einer Verletzung einer normativen Regel führt, so daß mögliche Verletzungen von vorneherein frühzeitig abgefangen werden. Die oben genannten aktiven Regeln dienen in diesem Falle nur als eine zusätzliche Rückfall- und Kontrollebene. Sollte tatsächlich einmal im automatischen Betrieb des Systems eine derartige Regel ausgelöst werden, weist dies auf einen Entwurfsfehler im System hin, der schnellstmöglich korrigiert werden sollte.

Bezüglich manueller Benutzerinteraktion des Systems werden diese aktiven Regeln zur Integritätsüberwachung allerdings auch im Regelbetrieb benötigt und eingesetzt. Fehlerhafte oder unzulässige Eingaben des Benutzers können auf diese Weise einfach zurückgewiesen werden.

3.4 Umsetzung der Modellierungstechniken in relationalen Datenbanksystemen

In den vorangegangenen Abschnitten wurden einige Modellierungstechniken vorgestellt, mit deren Hilfe das im Rahmen dieser Arbeit entwickelte System entworfen und beschrieben werden soll. Die Motivation für die Voraberstellung eines Modells wurde ausführlich dargelegt. Selbstverständlich ist das eigentliche Ziel, das modellierte System anschließend prototypisch zu implementieren und entsprechend ausführlich zu testen und entsprechend seinem Anwendungsgebiet einzusetzen.

Vorausschauend auf diesen Arbeitschritt sollen einige in Wissenschaft und Technik anerkannte Standardverfahren zur Umsetzung der Modelle in relationalen Datenbanksystemen kurz angerissen werden. Für Details zur Thematik (insbesondere die formalen Grundlagen) sei wiederum auf die entsprechende Fachliteratur, wie z. B. [Mai53], [Dat95] und [Kem97] verwiesen. An letzterer Quelle orientiert sich auch diese Darstellung.

Zunächst sollen ein paar einführende Worte zu Datenbanksystemen im allgemeinen und relationalen Systemen im speziellen erfolgen. Ein Datenbanksystem besteht aus einer in der Regel recht großen Menge an Daten sowie einem Satz von Programmen für einen effektiven und standardisierten Zugriff auf diese Daten. Die Daten werden auch als Datenbasis bezeichnet, die Menge der Programme ist ein sogenanntes Datenbankmanagementsystem (DBMS). Entsprechende Anwendungssoftware kann über wohldefinierte Schnittstellen (in der Regel mit Hilfe eine Abfragesprache) zum Datenzugriff mit dem DBMS kommunizieren, so daß man sich im Rahmen der Softwareentwicklung nicht mehr mit den Details der physischen Datenverwaltung beschäftigen muß (vgl. Abb. 3-9).

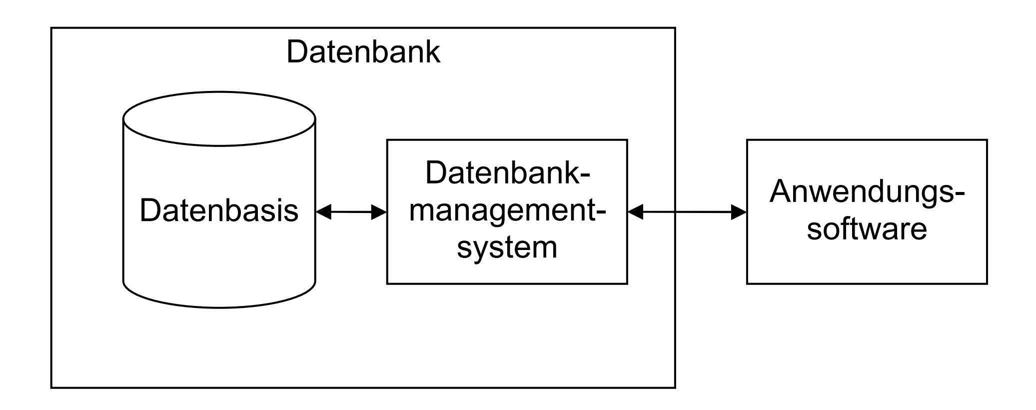


Abb. 3-9: Datenbanksystem

Wesentliche Aufgaben des DBMS sind die Vermeidung von Redundanzen und Inkonsistenzen in der Datenbasis, die Sicherstellung eines umfassenden Datenzugriffs, Ermöglichung eines Mehrbenutzerzugriffs, die Vermeidung von ungewolltem Datenverlust, die Sicherstellung der Datenintegrität, die Verwaltung von Zugriffsrechten auf die Daten und die bereits oben angedeutete Reduktion von Systementwicklungskosten durch Bereitstellung einer Schnittstelle.

Das Datenbanksystem basiert auf einem vorgegebenen Datenmodell, welches die „Infrastruktur“ für Datenmodellierung der realen Welt bereitstellt. Es legt die Modellierungskonstrukte fest, mit denen man ein Informationsabbild der realen Welt generieren kann. Damit kann einerseits eine Beschreibung der Datenobjekte und andererseits eine Festlegung der anwendbaren Operationen und deren Wirkungen festgelegt werden. Dies entspricht semantisch einer Programmiersprache, die in zwei Teilsprachen unterteilt werden kann, der Datendefinitionssprache (DDL) und der Datenmanipulationssprachen (DML). Die DDL dient dazu, Schemata („Typen“) für gleichartige Datenobjekte zu definieren. Die DML besteht aus der Anfragesprache und der „eigentlichen“ Manipulationssprache für Änderungen, Einfügungen und Löschungen von Daten. Die Anwendungssoftware greift in der Regel mittels DML-Kommandos auf die Datenbank zu.

Das heute am weitesten verbreitete Datenmodell ist das relationale Modell. Dieses hat einen relativ einfachen Grundaufbau. Es besteht überwiegend aus flachen Tabellen (Relationen), deren Zeilen den einzelnen Datenobjekten entsprechen. Die Relationen können über verschiedenen Operationen mengenorientiert verknüpft und verarbeitet werden.

Ein Relationenschema besteht entsprechend aus dem Namen der Relation und einem Tupel aus den Attributen (Spalten) dieser Relation, jeweils bestehend aus dem Attributsnamen und dem Wertebereich (Typ) des Attributs:

Relation1: { [<u>Attribut1: wertbereich1</u>, Attribut2: wertbereich2] }

Der Primärschlüssel einer Relation, der jeden Datensatz der Relation eindeutig beschreibt, wird durch Unterstreichung gekennzeichnet.

Ein Entity-Relationship-Modell läßt sich folgendermaßen in ein relationales Schema übertragen:

Entitytypen (ohne Generalisierung) lassen sich einfach in einer Relation ausdrücken, indem der Entityname zum Relationsnamen und die Entityattribute zu Relationenattributen werden. Die Schlüsselattribute können einfach aus dem ER-Modell übernommen werden.

Für Relationshiptypen ist das Verfahren etwas komplizierter: Zunächst einmal wird jedem Relationshiptyp im ER-Modell eine Relation im Relationenmodell zugeordnet. Der Name des Relationshiptyps wird wiederum zum Relationennamen. Die Attribute der Relation ergeben sich aus der Gesamtheit aller Primärschlüsselattribute der dem Relationshiptyp zugeordneten Entitytypen, ergänzt um alle dem Relationshiptyp direkt zugeordneten Attribute. Die Primärschlüsselattribute identifizieren das jeweils der Relationship zugeordnete Entity vollständig. Sie werden in diesem Zusammenhang als Fremdschlüssel bezeichnet. Wenn verschiedene an dem Relationshiptyp beteiligte Entitytypen die gleichen Attributbezeichner verwenden, kann eine Umbenennung der Attribute notwendig sein, um die Eindeutigkeit innerhalb der Relation sicherzustellen. Wiederum werden die Schlüsselattribute der Relation durch Unterstreichen gekennzeichnet. Bei einer N:M-Beziehung setzt sich dieser aus der Menge sämtlicher Fremdschlüsselattribute zusammen. Bei einer 1:N-Beziehung reicht der Fremdschlüssel des „N-fach“ auftretenden Entitytyps als Relationenschlüssel aus, da der Partner der Beziehung sich per Definition der Beziehung eindeutig von diesem Entity ableitet

(Funktionencharakter der Beziehung). Bei der Existenz von 1:1- bzw. 1:N-Beziehungen kann es daher sein, daß mehrere Relationen der Datenbank einen identischen Schlüssel haben. Dadurch können diese Relationen auch in einer einzigen Relation zusammengefaßt werden.

Wir bereits weiter oben angedeutet, enthält das relationale Datenmodell keine Vererbungskonstrukte. Hier gibt es zwei verschiedene Lösungen zur „Auflösung“ von Generalisierungen im ER-Modell. Eine Möglichkeit ist, die Oberklasse vollständig zu eliminieren und alle Attribute in die jeweiligen Unterklassen zu verschieben, die dann entsprechend in Relationen umgewandelt werden können. Die Alternative besteht darin, beide Klassen in Relationen umzuwandeln und den Primärschlüssel der Oberklasse ebenfalls als Primärschlüssel der Unterklassen zu verwenden (Begrenztes „Kopieren“ der Attribute), so daß der Zusammenhang der beiden Objekte (die eigentlich dasselbe reale Objekt darstellen) hergestellt werden kann.

Die Verknüpfung von Relationen kann mit Hilfe der Anfragesprachen erfolgen. Diese Anfragesprachen basieren formal allesamt auf der bereits in 3.3 erwähnten relationalen Algebra bzw. den ebenfalls dort bereits beschriebenen Relationenkalkülen. Diese Sprachen sind abgeschlossen, d. h. die Ergebnisse der Anfragen sind wiederum Relationen. Alle besitzen die gleiche Ausdruckskraft und sind daher gegenseitig austauschbar. Jede weitere Anfragesprache, die die Mächtigkeit der relationalen Algebra besitzt, wird als relational vollständig bezeichnet.

Da sowohl die relationale Algebra als auch die Kalküle nicht-intuitiven, eher formalen Charakter haben, wird im praktischen Einsatz in der Regel das deutlich benutzerfreundlichere und standardisierte SQL verwendet, das sich stärker an die an die englische Sprache orientiert. Dabei handelt es sich um eine auf dem Tupelkalkül basierende deklarative Datenbanksprache, die in den 1970er Jahren unter dem Namen SEQUEL [Cha74] bei IBM entwickelt wurde und 1986 von der ANSI bzw. 1987 von der ISO als Standard verabschiedet wurde. Der letzte Revision der ISO ist SQL:2008 [SQL08].

Alle bedeutenden auf dem Markt befindlichen Datenbanksysteme realisieren weitgehend die im Standard definierten Konzepte, wobei es teilweise Abweichungen in der Syntax geben kann.

Von besonderem Interesse im Rahmen dieser Arbeit ist dabei, daß SQL bereits standardmäßig Sprachkonstrukte bereitstellt, die die im vorigen Abschnitt vorgestellten deduktiven, aktiven und normativen Regelkonzepte relativ einfach mit Hilfe eines vorhandenen DBMS mit entsprechenden Inferenz- bzw. aktiver Komponenten implementierbar machen. Damit kann der Regelsatz im Modell des Systems mit geringem Aufwand in entsprechende SQL-Konstrukte umgewandelt werden, die über bestehende Schnittstellen des DBMS der Datenbank übermittelt werden können. Durch Nutzung dieser fertigen Komponenten eines geeigneten DBMS braucht sich der Systementwickler nicht mehr mit allgemeinen Implementierungsdetails der Regelkomponenten zu beschäftigen, sondern kann sich auf die eigentlichen Kernaufgaben der Systementwicklung konzentrieren.

An dieser Stelle eine vollständige Einführung in die Abfragesprache SQL zu geben, würde den Rahmen der Arbeit bei weitem sprengen, so daß der Fokus auf die Grundbausteine und die Implementierung der in 3.3 genannten Regelkonstrukte liegen soll.

Der elementare Anfrageausdruck in SQL besteht aus drei Komponenten, einem SELECT-Ausdruck, gefolgt von einem FROM-Ausdruck und (optional) abgeschlossen mit einem WHERE-Ausdruck. Der SELECT-Ausdruck gibt hierbei die Attribute der Ergebnisrelation wieder, der FROM-Ausdruck die benötigten vorhandenen Relationen und der WHERE-Ausdruck schränkt die Resultate um entsprechende Auswahlbedingungen ein.

Eine einfache Verknüpfung zweier Relationen ist daher folgendermaßen möglich:

```
SELECT R1.Attr1, R2.Attr1
FROM R1, R2
WHERE R1.Attr2 = R2.Attr2 AND R1.Attr1 > 50;
```

Hierbei wird das Attribut Attr1 der Relation R1 und das Attribut Attr1 der Relation R2 in einer neuen Relation ausgegeben. Hierbei muß für jeden neuen Datensatz gelten, daß die mit den beiden ausgegebenen Attributen jeweils im selben Datensatz befindliche Attribute Attr2 der Ausgangsrelationen R1 und R2 identisch sein müssen und der ausgegebene Attr1 Wert von R1 größer als 50 sein muß.

Ein analoger Ausdruck im „Datalog-Stil“ wäre entsprechend:

```
zielrelation(A,B) ← r1(A,X), r2(B,X), A > 50.
```

Hier wird bereits die Verwandtschaft zu den deduktiven Regeln offensichtlich, die ja schließlich auch nichts anderes als Kalkülausdrücke darstellen.

Deduktiven Regeln werden in relationalen Datenbanken als Sichten realisiert. Eine Sicht ist in diesem Zusammenhang eine Beschreibung einer für eine bestimmte Benutzergruppe interessante Datenmenge. Damit können „virtuelle“ Relationen erzeugt werden, die von DBMS analog zu echten Tabellen behandelt werden können, also auch in weiteren Abfragen als eine Art „Makro“ verwendet werden können.

Das entsprechende Sprachkonstrukt in SQL ist der View. Die View-Syntax in SQL leitet sich dementsprechend auch aus dem folgenden DDL-Teil der Sprache zur Definition neuer Tabellen ab, für das hier ein Beispiel angegeben wird:

```
CREATE TABLE Relationenname
( Attr1 integer not null,
  Attr2 varchar(5) );
```

Hier wird eine Relation mit Namen Relationenname angelegt, die aus zwei Attributen besteht. Attr1 ist eine Zahl und darf keinen leeren (NULL) Wert enthalten. Attr2 ist eine Zeichenkette aus maximal fünf Zeichen.

Die vollständige SQL-View-Syntax ist entsprechend dem obigen Beispiel:

```
CREATE VIEW Zielrelation(Attr1, Attr2) AS
  SELECT R1.Attr1, R2.Attr1
  FROM R1, R2
  WHERE R1.Attr2 = R2.Attr2 AND R1.Attr1 > 50;
```

Wegen der ansonsten fehlenden Eindeutigkeit wurde Attr1 aus R2 innerhalb der Zielrelation in Attr2 umbenannt.

Normative Regeln werden im Kontext relationaler Datenbanken als Integritätsbedingungen bzw. Constraints realisiert. Das DBMS wacht automatisch darüber, daß die Datenbank in einem konsistenten Zustand bzgl. aller Relationen sämtliche deklarierten Constraints erfüllt. Die Integritätsbedingungen können entweder explizit definiert werden, durch die Schemadefinition implizit erzwungen werden oder dem Datenmodell inhärent sein.

Inhärente Bedingungen leiten sich z. B. aus den vorgegebenen Wertebereichen für Attribute ab. Ein impliziter Constraint ist beispielsweise die „not null"-Angabe in der obigen „create table" Definition.

Die expliziten Integritätsbedingungen werden am Ende der bereits erwähnten CREATE TABLE-Syntax nach der Auflistung der einzelnen Attribute definiert. Hier ein Beispiel mit verschiedenen Typen von Integritätsbedingungen:

```
CREATE TABLE Relationenname
( …,   -- Definition der Attribute
PRIMARY KEY (Attr1, Attr2),
FOREIGN KEY (Attr3) REFERENCES R1(Attr1),
CHECK Attr4 in ('A','B','C'),
CHECK Attr5 BETWEEN 2 AND 5);
```

Die Kombination der Attribute Attr1 und Attr2 stellt einen Primärschlüssel dar, der jeden Datensatz in Relationenname eindeutig definiert. Duplikate sind daher unzulässig. Jeder Wert von Attr3 muß, solange er nicht NULL ist, einem Schlüsselwert Attr1 aus der Relation R1 entsprechen. Dieses wird auch als referentielle Integrität bezeichnet. Der zulässige Wertebereich von Attr4 wird auf die Werte ‚A', ‚B' und ‚C' begrenzt. Ebenso wird der zulässige Wertbereich von Attr5 auf einen Zahlenwert zwischen 2 und 5 begrenzt.

Komplexere Constraints können ferner tabellenübergreifend mit Hilfe sogenannter Assertions explizit definiert werden.

Schließlich verbleiben als letzter Regeltyp noch die aktiven Regeln, die im relationalen Datenbankkontext als Trigger realisiert werden. Dabei handelt es sich um eine benutzerdefinierte Prozedur, die bei Erfüllung einer bestimmten Bedingung automatisch vom DBMS gestartet wird. Sie werden primär zur Integritätserhaltung und für Berechnungen eingesetzt.

SQL-Trigger bestehen aus vier Teilen: Einer CREATE TRIGGER-Anweisung, gefolgt von einem Namen, einer Definition des Auslösers (Event), einer einschränkenden Bedingung (Condition) und der Reaktions-Prozedurdefinition (Action).

Ein Beispiel ist:

```
CREATE TRIGGER r1_after_delete
AFTER DELETE ON r1
FOR EACH ROW
WHEN EXISTS (SELECT * FROM r2
WHERE r2.attr2 = :old.attr1)
BEGIN
  DELETE FROM r2
  WHERE r2.attr2 = :old.attr1
END;
```

Im Trigger namens r1_after_delete geht es um die Sicherstellung der referentiellen Integrität bei Löschung eines Datensatzes aus der Relation r1. Für jede Löschung wird geprüft, ob es in der Relation r2 Datensätze gibt, die in deren Attribut attr2 auf den entsprechend gelöschten Primärschlüssel aus r1 verwiesen wird. Falls dies der Fall ist, sollen diese Datensätze ebenso aus r2 gelöscht werden.

4 Konzeptuelle Modellierung des Eisenbahnbetriebs

Bevor das im Rahmen dieser Arbeit entwickelte Prioritätenvergabe- und -optimierungssystem genauer in seinen statischen, dynamischen und funktionalen internen Aspekten spezifiziert bzw. modelliert werden kann, soll zunächst einmal die bereits existierende Anwendungsumgebung eines solchen Systems, d. h. der mit Hilfe von Prioritäten disponierte Eisenbahnbetrieb, mit Hilfe der im vorigen Kapitel eingeführten Modellierungstechniken im dafür notwendigen Detaillierungsgrad („so detailliert wie nötig, so einfach wie möglich“) konzeptuell beschrieben werden. Wie dort bereits erläutert wurde, soll sich die Modellierung auf die im Zusammenhang mit Prioritätenvergabe und -optimierung relevante Aspekte des Eisenbahnbetriebs beschränken, um darauf aufbauend das entsprechende System entwickeln und darstellen zu können. Die Modellierung des Anwendungsgebietes stellt den „Status quo“ dar und dient unter anderen zur Spezifizierung der Input- und Output-Schnittstellen des im weiteren Verlauf dieser Arbeit entwickelten Prioritätenvergabe- und -optimierungssystems. Allerdings wäre die Modellierung zu eng gefaßt, wenn sich ausschließlich unter diesem Aspekt erfolgen würde. Das Modell sollte sämtliche Aspekte des Betriebes erfassen, die sachlich im engen Zusammenhang mit den Aufgaben des geplanten Systems stehen, unabhängig von der Spezifikation oder gar Implementierung desselben. Auf der anderen Seite sollten sich aber alle externen Schnittstellen des Systems im hier aufgezeigten Modell auch wiederfinden.

Da der Prioritätenvergabe- und -optimierungsprozeß unabhängig von der Umsetzung konkreter Dispositionsmaßnahmen oder sonstiger Aspekte der Betriebsdurchführung erfolgen soll, kann der Schwerpunkt der Modellierung auf statische Aspekte gelegt werden. Der konkrete Betriebsablauf inkl. der Konflikterkennung und prioritätenbasierten Konfliktlösung durch den Disponenten müssen nicht in der gesamten Breite berücksichtigt werden. Wirklich für die Prioritätenvergabe wesentliche dynamische Aspekte der Anwendungsumgebung können in dem benötigten Umfang einfach mit Hilfe von diskreten Ereignissen (Ereignisobjekten) ebenfalls in einem statischen Objektmodell mit berücksichtigt werden. Dabei werden „Schnappschüsse“ von für die Prioritätenvergabe bedeutsamen Aspekten des Anwendungsgebietes zu einem konkreten Zeitpunkt gemacht und als ein Ereignis „codiert“, das später über eine Schnittstelle nach „außen“ übermittelt werden kann.

Die gesamte Anwendungsumgebung ist im erforderlichen Detaillierungsgrad nicht überschaubar. Daher ist es für den eigentlichen Zweck der Modellierung nicht sinnvoll, alles in einem ER-Diagramm darzustellen. Es ist vielmehr empfehlenswert, diese in verschiedene Sichten aufzuteilen. Für die jeweilige Sicht wird ein spezieller Entwurf durchgeführt, der sich auf den notwendigen Bedarf beschränkt. Aufgrund der jeweiligen Anforderungen sind die verschiedenen Sichten des Entwurfes in der Regel nicht disjunkt, sondern überschneiden sich in mehr oder weniger großem Umfang. Sie können anschließend zu einem globalen Schema zusammengefaßt werden, welches redundanz- und widerspruchsfrei ist und wo die Synonyme und Homonyme bereinigt sind.

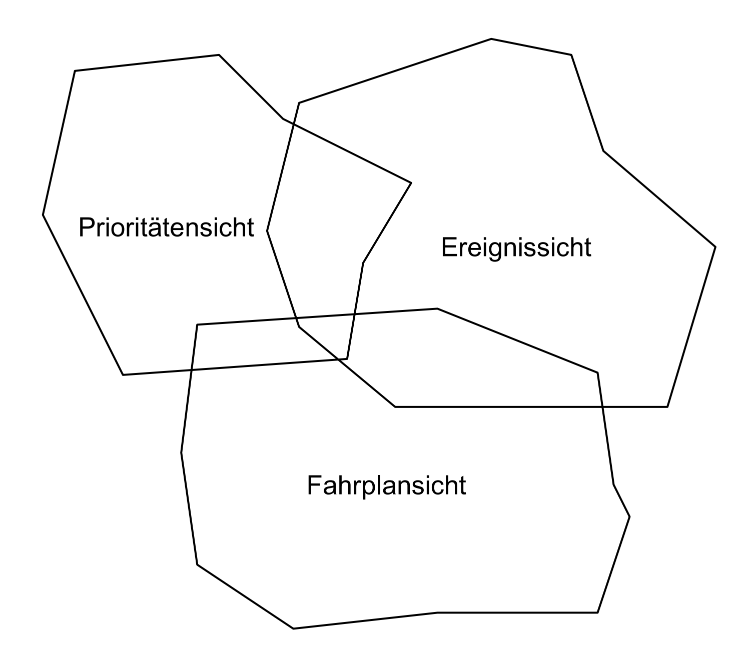


Abb. 4-1: Anwendungsumgebung des Prioritätenvergabe- und -optimierungssystems

Für die prioritätenbasierte Eisenbahnbetriebsdisposition bieten sich drei zentrale Sichten an, nämlich die Fahrplansicht, die Ereignissicht und die Prioritätensicht (vgl. Abb. 4-1). Diese Aufteilung in drei Sichten erklärt sich einfach aus der geplanten Funktionalität und den Schnittstellen eines zukünftigen Prioritätenvergabe- und -optimierungssystems. Die Prioritätensicht steht im Zusammenhang mit dem „Output" des Systems in seine Anwendungsumgebung, insbesondere einem Dispositionssystem zur Konfliktlösung. Die Ereignissicht entspricht dem „Input" aus der Anwendungsumgebung in das System. Aufgrund der übermittelten Ereignisse schlägt das System für die Konfliktlösung im Rahmen der Disposition zu verwendende Prioritäten vor. Die Fahrplansicht vereinigt die Eisenbahninfrastruktur, die fahrenden Züge und die temporäre Zuordnung der Züge auf diese Infrastruktur. Dieser Bereich bildet damit sozusagen eine Grundlage, die eine Einordnung der beiden zuvor genannten Sichten in den Gesamtkontext ermöglicht.

Im folgenden sollen die drei Sichten genauer betrachtet und mit Hilfe von ER-Diagrammen modelliert werden.

4.1 Fahrplansicht

In der Fahrplansicht werden grundlegende Daten in einem im Zusammenhang mit Prioritätenvergabe und -optimierung relevanten Detaillierungsgrad modelliert. Hierzu kann die Fahrplansicht in drei thematische Bereiche bzw. „Untersichten" gruppiert werden, nämlich die Eisenbahninfrastruktur, die Züge und als Bindeglied zwischen diesen Bereichen der eigentliche Fahrplan (vgl. Abb. 4-2). Der Fahrplan selbst kann noch einmal unterschieden werden in einen langfristig gültigen Jahresfahrplan und einen jeweils aktuellen Tagesfahrplan.

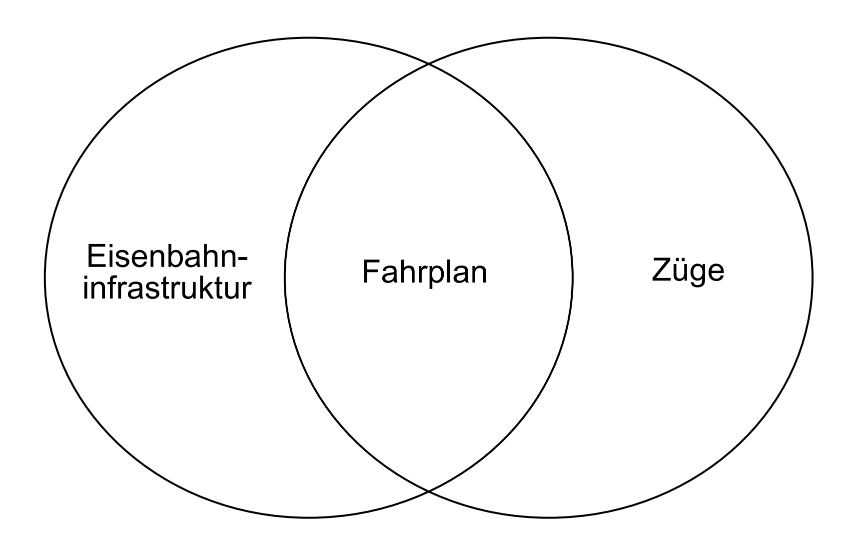


Abb. 4-2: „Untersichten" der Fahrplansicht

4.1.1 Eisenbahninfrastruktur

Insbesondere bei der Eisenbahninfrastruktur ist die Frage des Detaillierungsgrades von entscheidender Bedeutung. Für Dispositionssysteme und damit verbundene exakten Fahrtzeitenrechungen wird z. B. eine weichen- und signalgenaue Spurplan-Infrastrukturmodellierung benötigt. Im Zusammenhang mit Prioritäten sind derartige Details allerdings nicht von Interesse – Zugprioritäten sollen sich nicht an einzelnen Weichen und Signalen verändern –, so daß auf ein mikroskopisches Modell verzichtet werden kann. Informationen über die Eisenbahninfrastruktur wird einerseits für die Vergabe von Zugprioritäten auf Strecken benötigt und andererseits für die Vergabe von Anschlußprioritäten in Bahnhöfen. Zusätzlich wird sie für den Kontext von relevanten betrieblichen Ereignissen benötigt.

Daraus ergeben sich die folgenden Vorgaben für das Infrastrukturmodell: Für Bahnhöfe wird ein halteplatzgenaues Modell benötigt, um Anschlußübergänge von Reisenden und Güterwagen geeignet im Zusammenhang mit Mindestübergangszeiten untersuchen zu können, die für Anschlußprioritäten von Interesse sind. Im Güterverkehr können auch mehrere Halteplätze mit identischen Mindestübergangszeiten (z. B. Einfahr- und Ausfahrgruppen von Rangierbahnhöfen) zusammengefaßt werden

Die Verbindungen zwischen Bahnhöfen sollen durch Streckenabschnitte erfolgen. Dabei soll jedes durchgehende Hauptgleis zwischen zwei modellierten Bahnhöfen als eigenständiger Streckenabschnitt modelliert werden. Dies erlaubt, parallele Gleise prioritätentechnisch unterschiedlich behandeln zu können, z. B. bestimmte Verkehrsarten auf bestimmten Gleisen zu bevorzugen. Mit Ausnahme von zum Übergang zwischen verschiedenen Streckenabschnitten benötigten Überleit- und Abzweigstellen brauchen, wie bereits oben erwähnt, weitere Details für die beabsichtigte Anwendung nicht modelliert zu werden.

Daraus ergibt sich eine Modellierung der Infrastruktur entsprechend eines Hypergraphen. Die Graphknoten repräsentieren die einzelnen Bahnhöfe, Abzweig- und Überleitstellen. Die Halteplätze der Bahnhöfe entsprechen Unterknoten des jeweiligen Knoten, die durch Rei-

senden- bzw. Güterwagenübergänge repräsentierende Unterkanten verbunden sind. Die Bahnhöfe sind untereinander entsprechend durch Streckenabschnittskanten verbunden, wobei die Kanten sowohl gerichtet sein können (falls das entsprechende Hauptgleis nur in einer Richtung befahrbar ist) als auch ungerichtet sein können (falls Verkehr in beide Richtungen möglich ist, insbesondere im Zusammenhang mit eingleisigen Strecken). Die Eisenbahninfrastruktur wird als unverändert angenommen.

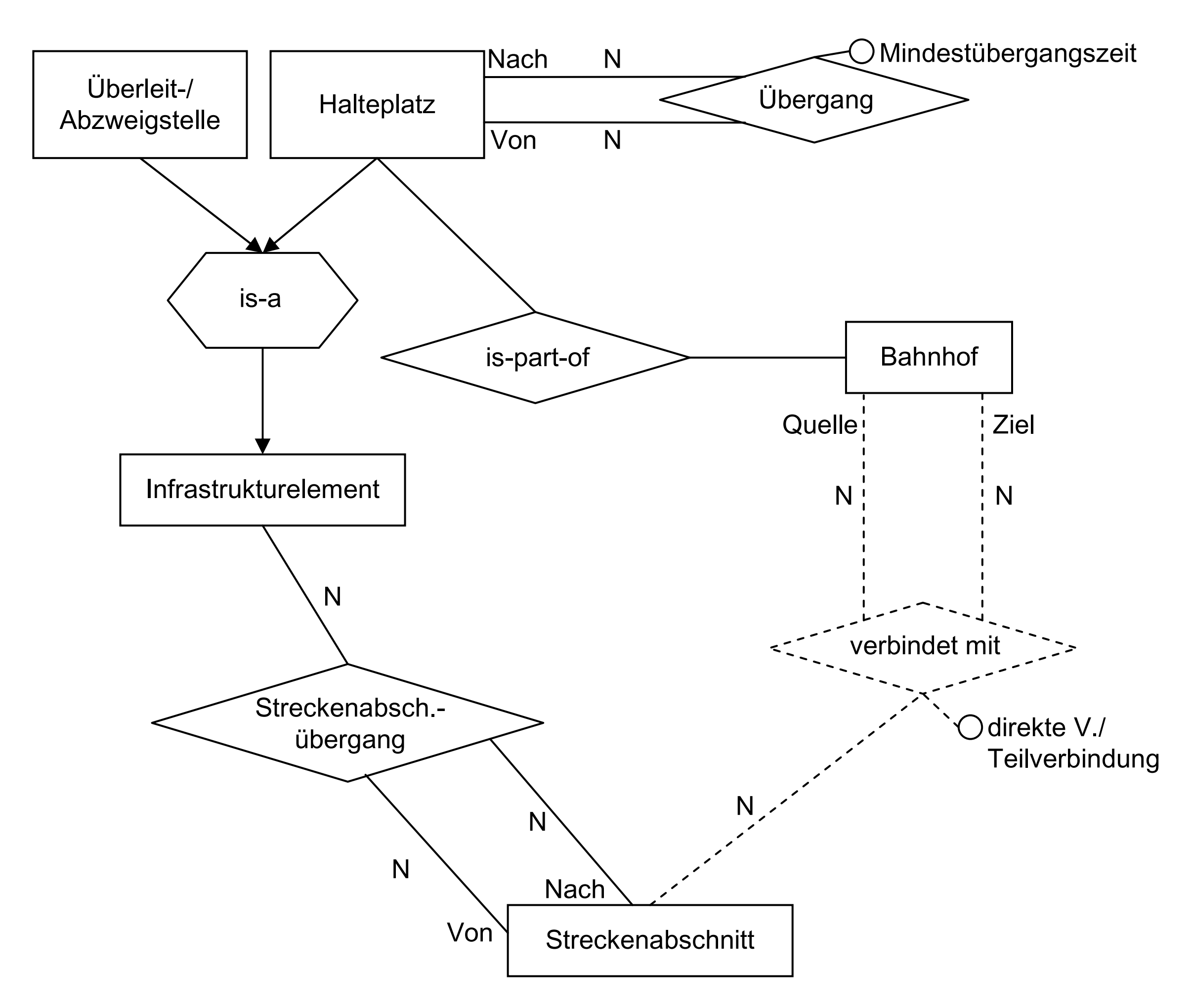


Abb. 4-3: ER-Diagramm Eisenbahninfrastruktur

Eine Modellierung mit einem ER-Diagramm ist Abb. 4-3 zu entnehmen. Man beachte dabei, daß die Verbindungen zwischen Bahnhofsknoten hier abgeleitet werden aus Streckenübergangsinformationen innerhalb der Halteplätze, Überleit- und Anschlußstellen. Dazu dient der Relationshiptyp „Streckenabschnittsübergang“, der in ein Infrastrukturelement eingehendende Streckenabschnitte mit aus dem Infrastrukturelement ausgehenden Streckenabschnitten verknüpft.

4.1.2 Züge

Züge stellen die mobilen Objekte dar, die auf der soeben eingeführten Eisenbahninfrastruktur verkehren sollen. Analog zur Eisenbahninfrastruktur kann im Zusammenhang mit Festlegung von Prioritäten wiederum auf technische Details verzichtet werden, die z. B. für aus der

Prioritätensicht unbedeutende Fahrtzeitenrechnungen benötigt werden. Auch einzelne Triebfahrzeuge bzw. Waggons brauchen nicht exakt modelliert werden, da der Zug für das im Kontext von Prioritäten als atomare Einheit fungiert.

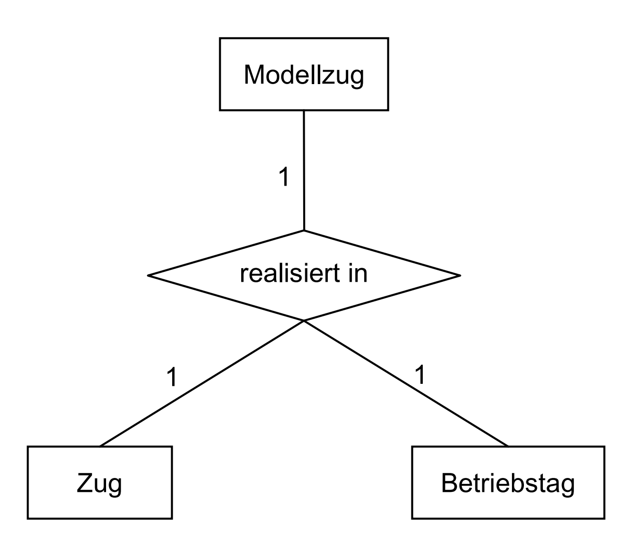


Abb. 4-4: ER-Diagramm Züge

Als zentrales Modellierungsobjekt soll in diesem Zusammenhang der „Modellzug" dienen. Dieser ist ein bestimmtes Zugkonzept im Jahresfahrplan, das Kunden entsprechend kommuniziert werden kann. An allen seinen Betriebstagen wird der Modellzug durch einen konkreten Zug realisiert. In Abb. 4-4 findet sich das entsprechende ER-Diagramm.

4.1.3 Jahresfahrplan

Der Jahresfahrplan ist das Bindeglied zwischen Eisenbahninfrastruktur und Modellzügen und stellt somit das eigentliche Modellierungsziel der Fahrplansicht dar. Für den Jahresfahrplan gilt natürlich auch wieder das Grundprinzip, sich auf im Zusammenhang mit Prioritätenverwaltung- und -optimierung relevante Aspekte zu beschränken. Daher muß man keine genauen Bildfahrpläne o. ä. verwalten, sondern kann sich auf die Vorgänge in den Bahnhöfen konzentrieren.

Dazu gehören standardmäßige Ankunfts- und Abfahrtszeiten der Modellzüge im Bahnhof und die Definition von geplanten Anschlüssen zwischen zwei Modellzügen. Dies steht im Zusammenhang mit den zwei zentralen und mit Hilfe der Prioritätenvergabe zu optimierenden Aspekte der Eisenbahnbetriebsdisposition, nämlich Pünktlichkeit und Anschlußsicherheit. Schließlich sollen noch die Modellzuglaufwege über die Streckenabschnitte berücksichtigt werden. Für die Verkehrshalte sind Bonus-Malus-Funktionen bzgl. der verkehrswirtschaftlichen Bewertung von Ankunftsverspätungen vorzusehen. Die Bewertungsfunktion kann abhängig vom Verspätungsbetrag (i. d. R. in Minuten) sein. Aus den Diskussionen ergibt sich für Standardlaufwege und Verkehrshalte das ER-Diagramm in Abb. 4-5.

Zu diesem ER-Diagramm kann man sicherlich noch anmerken, daß es natürlich auch möglich ist, den Entitytyp Verkehrshalt mit dem Relationshiptyp Modellzughaltliste zusammenzufassen. Da aber in der weiteren Modellierung der Verkehrshalt eine zentrale Rolle insbeson-

dere im Zusammenhang mit Verspätungen spielt, ist eine Modellierung als Entitytyp an dieser Stelle vorzuziehen.

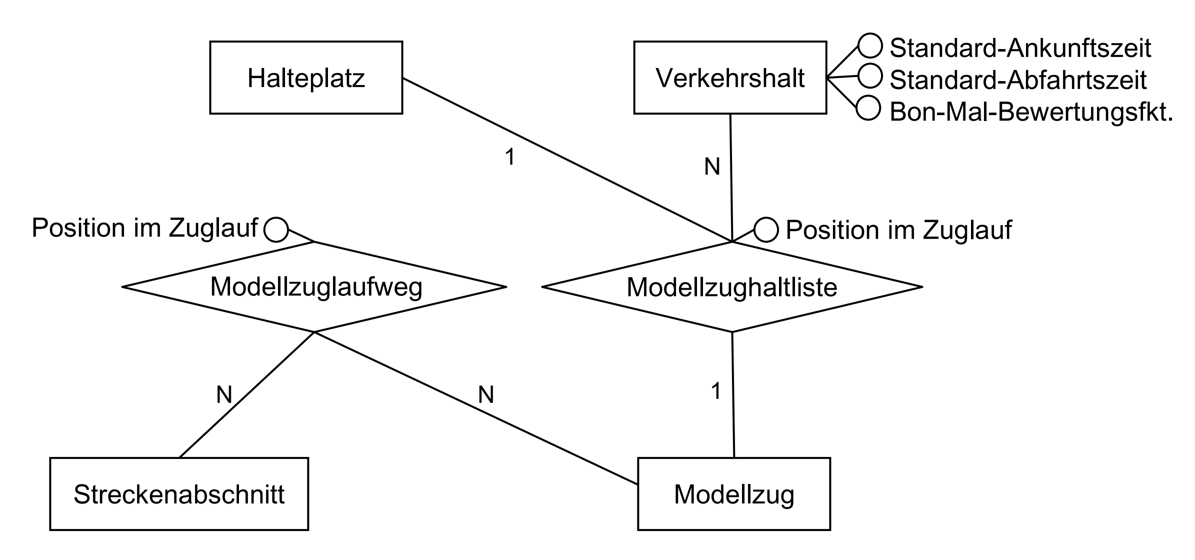


Abb. 4-5: ER-Diagramm Modellzuglaufwege und -verkehrshalte

Für die Behandlung von fahrplanmäßigen Anschlüssen wird analog das ER-Schema in Abb. 4-6 vorgeschlagen. Auch hier wird wiederum ein Anschluß basierend auf Modellzugpaaren modelliert. Bzgl. der Modellierung des Anschlusses als Entitytyp gilt sinngemäß dasselbe, was bereits zu Verkehrshalten gesagt wurde. Man beachte, daß sich die zeitbezogenen Attribute eines Anschlußobjektes einfach aus den Mindestübergangszeiten der planmäßig benutzten Halteplätze sowie den Ankunfts- und Abfahrtszeiten der Züge ableiten lassen. Das Attribut „verkehrlich/technisch“ unterscheidet zwischen verkehrlichen Anschlüssen zwischen zwei Modellzügen für Reisenden- und Frachtgutübergänge und Umlaufverknüpfungen zum Übergang von Wagenmaterial. Für die Bonus-Malus-Bewertung kann ein Pauschalbetrag für Anschlußbrüche festgelegt werden.

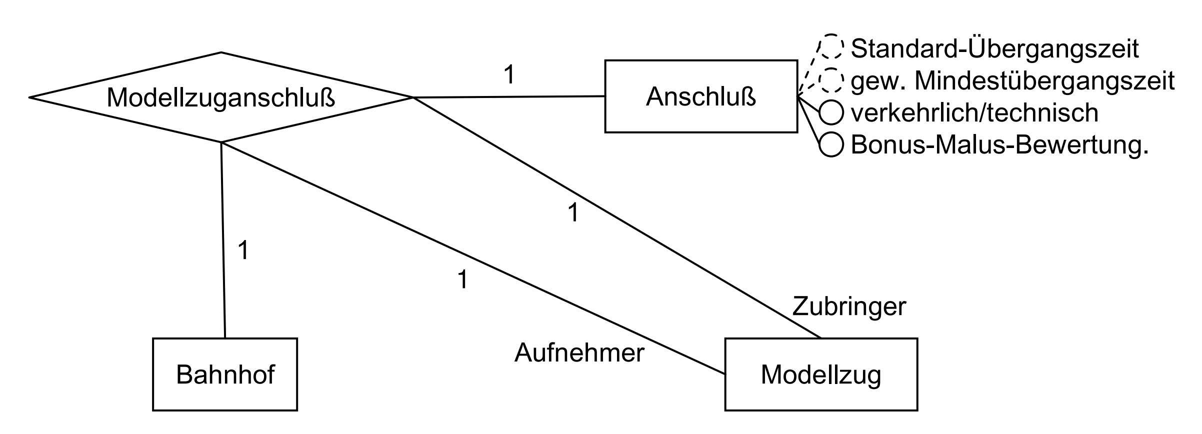


Abb. 4-6: ER-Diagramm Modellzuganschlüsse

4.1.4 Tagesfahrplan

Grundsätzlich entspricht der Tagesfahrplan von Zügen dem jeweiligen Jahresfahrplan der zugeordneten Modellzüge. Im Detail kann es allerdings von Betriebstag zu Betriebstag zu Abweichungen bzgl. Verkehrszeiten, Halteplätzen und Laufwegen kommen, insbesondere

im Zusammenhang mit Baustellen, längerfristigen Gleissperrungen oder technischen Problemen.

Dazu werden auf dem Entitytyp Zug basierende Relationshiptypen Zugaltliste, Zuglaufweg und Zuganschluß definiert (siehe Abb. 4-7 und 4-8).

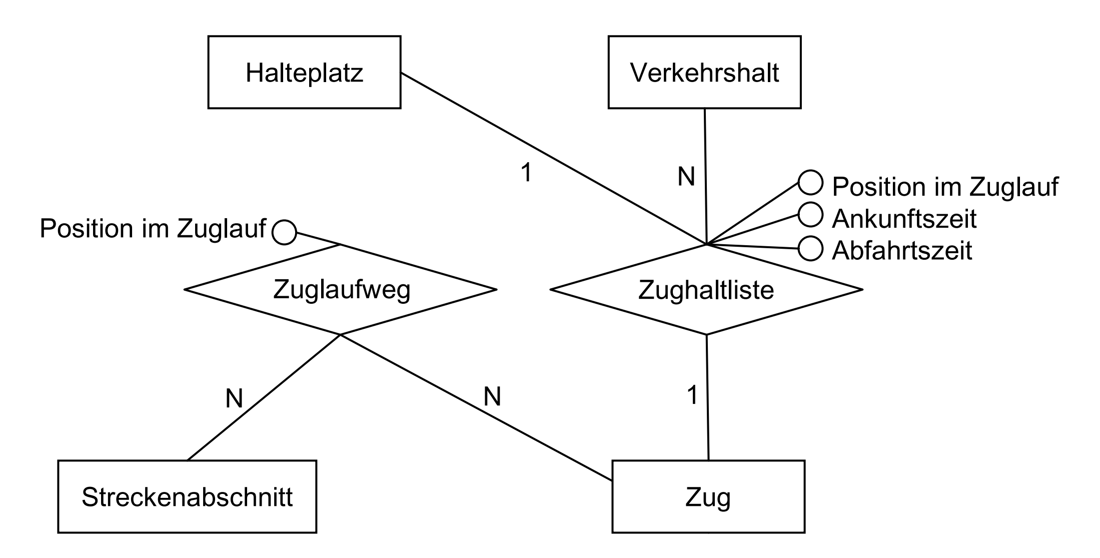


Abb. 4-7: ER-Diagramm Zuglaufwege und -verkehrshalte

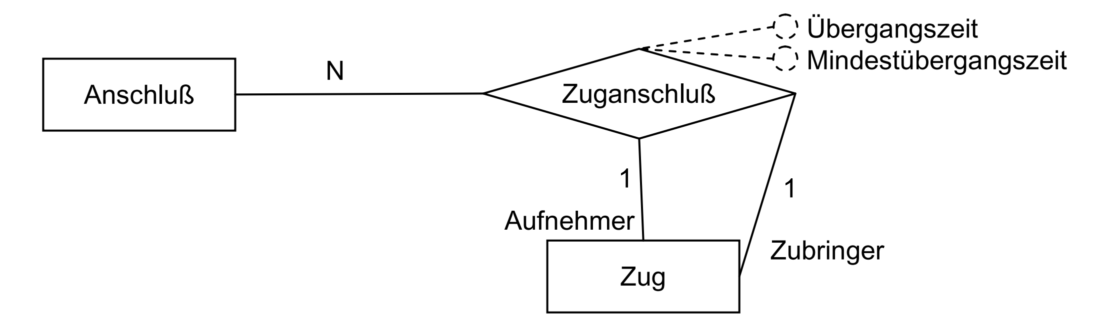


Abb. 4-8: ER-Diagramm Zuganschlüsse

4.2 Prioritätensicht

Prioritäten dienen im Rahmen der Disposition als Hilfe zur Aufstellung von Dispositionsmaßnahmen im Falle von erkannten Konflikten (siehe auch Ereignissicht in 4.3). Die Prioritätensicht entspricht dem Output des zu entwickelnden Prioritätenvergabe- und -optimierungssystems in seine Anwendungsumgebung. Prioritäten im Sinne des Anwendungsgebietes sind Dispositionsprioritäten, die in „Abschnittsweise Zugprioritäten" und „Anschlußprioritäten" unterteilt werden können. Die beiden Prioritätenarten werden im folgenden detailliert betrachtet.

4.2.1 Abschnittsweise Zugprioritäten

Abschnittsweise Zugprioritäten geben Rangstufen von Modellzügen auf bestimmten Prioritätenabschnitten an. Bei einem Prioritätenabschnitt handelt es sich um unter Umständen meh-

rere zusammenhängende Streckenabschnitte, innerhalb derer es keine Änderung der Rangstufen von Zügen geben soll. Um einen Prioritätenwechsel praktisch in der Disposition durchführen zu können, ist es unter Umständen notwendig, bzgl. Priorität „heruntergestufte" Züge an den Prioritätenwechselpunkten warten zu lassen, um relativ hochrangigen Zügen die Vorfahrt zu ermöglichen. Daher sollen nur solche Stationen für Übergänge zwischen Prioritätenabschnitten in Frage kommen, bei denen ausreichende Puffermöglichkeiten für Züge bestehen, um notwendige Warte- und Überholvorgänge für einen Prioritätentausch durchführen zu können. Ebenso sollte an Überleit- und Abzweigstellen für im Regelbetrieb befahrende Streckenabschnittsübergänge grundsätzlich auf Prioritätenabschnittswechsel verzichtet werden, da diese keine solchen Puffermöglichkeiten bieten und daher erhebliche betriebliche Behinderungen resultieren können.

Durch die Existenz der Pufferkapazitäten an der Prioritätenabschnittsgrenze ist eine Umsetzung der Abschnittsweisen Zugprioritäten im Rahmen der Disposition gewährleistet. Prioritätenabschnitte sollten dabei nicht zu klein sein, damit es im Verlauf eines Zuglaufes nicht zu dem Endkunden schwer vermittelbaren übermäßigen Prioritätenschwankungen kommt. Sie sollten aber auch nicht zu groß sein, um eine gewisse Flexibilität bei der Disposition beibehalten zu können.

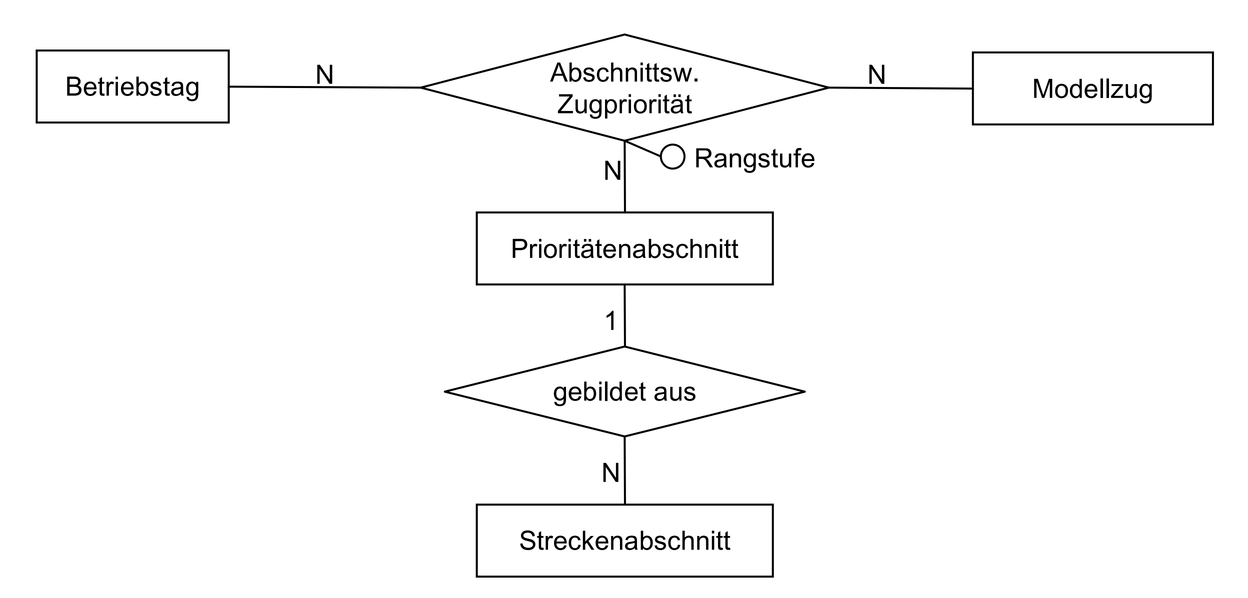


Abb. 4-9: ER-Diagramm Abschnittsweise Zugprioritäten

Die Zugprioritäten sind eine Hilfe für den Disponenten im Falle von Belegungskonflikten. Der im Konflikt befindliche Zug mit höherem Rang (entspricht einer niedrigen Nummer) erhält den Vorrang, der niederrangige Zug (entspricht einer höheren Nummer) muß warten. Zugprioritäten werden bzgl. eines Modellzugs an einem oder mehreren Betriebstagen für Prioritätenabschnitte vergeben. Durch die Bildung von Prioritätenabschnitten und Vergabe von Abschnittsweisen Zugprioritäten können implizit auch Vorrangstrecken für bestimmte Verkehrsarten (ggf. auch nur für bestimme zeitliche Abschnitte, z. B. Güterverkehr in der Nacht) definiert und unterschiedliche Ränge für den Gesamtlaufweg eines Zuges vergeben werden.

Da es sich bei Abschnittsweisen Zugprioritäten nicht wirklich um Objekte mit eigener Identität handelt, sondern dieses Konzept mehr Attributcharakter hat, wird die Zugpriorität nicht

als Entitytyp, sondern als Relationshiptyp modelliert. Es ergibt sich das Diagramm in Abb. 4-9.

4.2.2 Anschlußprioritäten

Bzgl. der Anschlußprioritäten gilt ähnliches wie für die Abschnittsweisen Zugprioritäten. Mit Hilfe der Anschlußprioritäten werden maximal zulässige Wartezeiten für den anschlußaufnehmenden Zug (über den Planabfahrtszeitpunkt hinaus) festgelegt. Aus der verspäteten Ankunftszeit des Anschlußzubringerzuges, der Mindestübergangszeit und der maximal zulässigen Wartezeit kann in der Disposition einfach entschieden werden, ob ein Anschluß eingehalten werden soll oder ob der aufnehmende Zug – ohne den Anschluß abzuwarten – pünktlich abfahren soll. Für Umlaufübergänge des Wagenparks von einem Zug auf einen anderen wird diese Regelung analog angewendet.

Aufgrund der obigen Definition von Prioritätenabschnitten als Abschnitte, innerhalb derer keine längeren Warte- und Überholvorgänge an Stationen durchgeführt werden können, kann ebenso angenommen werden, daß innerhalb eines Prioritätenabschnittes keine Anschlußprioritäten festgelegt werden können. Anschlüsse werden entsprechend bei Übergang von einem Prioritätenabschnitt in einen anderen Prioritätenabschnitt durchgeführt.

Auch die Anschlußpriorität wird aus dem gleichen Grund wie die Streckenpriorität nicht als Entitytyp, sondern als Relationshiptyp modelliert, woraus sich das ER-Schema für Anschlußprioritäten in Abb. 4-10 ergibt.

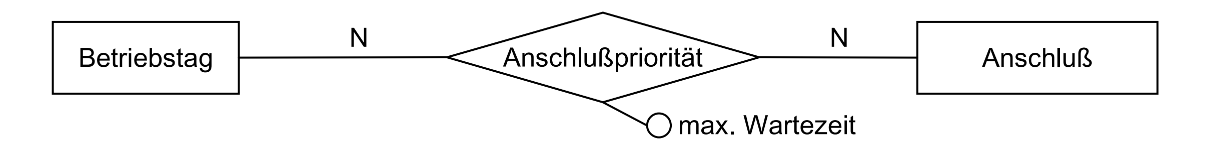


Abb. 4-10: ER-Diagramm Anschlußprioritäten

4.3 Ereignissicht

Ereignisse stellen den Input für das zu entwickelnde Prioritätenvergabe- und -optimierungssystem aus der Anwendungsumgebung dar. Hierbei stehen insbesondere die betriebliche Aspekte und die Auswertung von Bonus-Malus-Zahlungen im Vordergrund. Ereignisse bezeichnen ein Geschehen zu einem wohldefinierten und eindeutigen Zeitpunkt. Mit Hilfe der Ereignisse werden einzelne für die Prioritätenvergabe bedeutsame Schnappschüsse der dynamischen Aspekte der Anwendungsumgebung codiert. Damit kann die Modellierung des dynamischen Verhaltens des Eisenbahnbetriebes auf ein Minimum reduziert werden und für die Prioritätenermittlung nicht relevante interne Abläufe können entsprechend vernachlässigt werden.

Jedes signalisierte Ereignisobjekt hat einen konkreten Ereignistyp. Dabei ist es Teil der Modellierungsaufgabe, die Anzahl der Ereignistypen möglichst gering zu halten und sich auf

Zusammenhang mit Prioritäten relevante Aspekte zu beschränken. Man kann die dazu benötigten Ereignisse unterteilen in Betriebsereignisse und Bonus-Malus-Zahlungen.

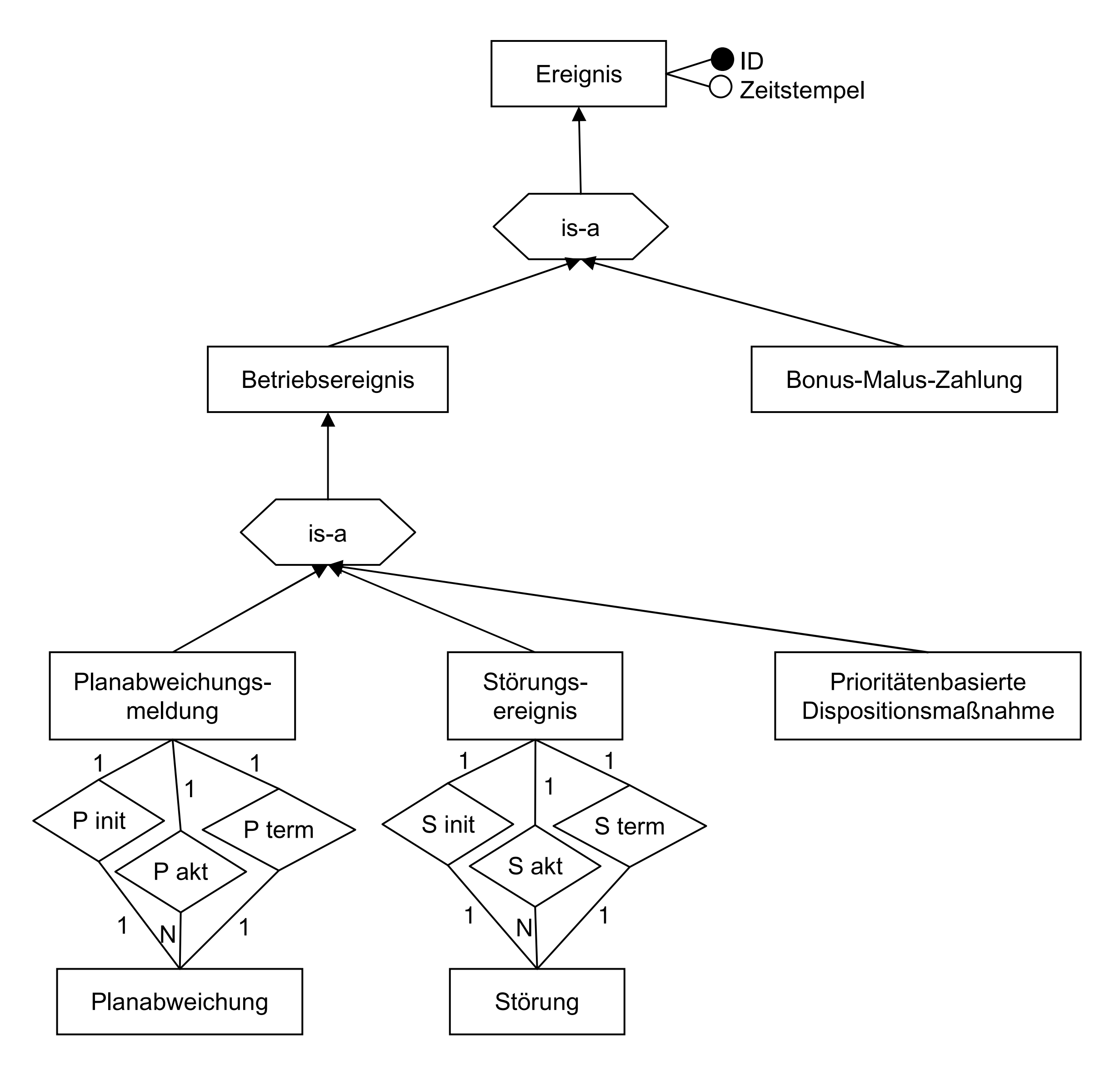


Abb. 4-11: ER-Diagramm Ereignisse

Als relevante Betriebsereignisse kommen Planabweichungsmeldungen (Verspätungen, Anschlußverluste usw.), Störungsereignisse bzgl. Infrastruktur und Fahrzeugen, und prioritätenbasierte Dispositionsmaßnahmen (insbesondere Halt- und Fahrtzeitenverlängerungen) in Betracht. Nicht prioritätenbasierte Dispositionsmaßnahmen, wie z. B. größere Umleitungen oder Zugausfälle sind in aller Regel direkt an eine Störung gekoppelt und können außerdem im Rahmen des Tagesfahrplan berücksichtigt werden, so daß eine eigene Modellierung nicht unbedingt notwendig ist. Planabweichungsmeldungen und Störungsereignisse können eine Planabweichung bzw. Störung initialisieren, aktualisieren oder terminieren. Bonus-Malus-Zahlungen dienen der Prioritätenevaluierung. Die Betriebsereignisse werden in der Regel von einem Dispositionssystem erkannt und übermittelt, für die Bonus-Malus-Zahlungen hingegen ist eine Qualitätsüberwachungskomponente verantwortlich, welche die Bonus-Malus-Zahlungen entsprechend dem vorgegebenen Bonus-Malus-System überwacht.

Die für das System relevanten Ereignisklassen im Verlauf der Betriebsdisposition können in einem Objektmodell mit Hilfe von ER-Diagrammen illustriert werden (Abb. 4-11).

Im folgenden sollen die Planabweichungen, Störungen und Prioritätenbasierten Dispositionsmaßnahmen, sowie die Bonus-Malus-Zahlungen genauer diskutiert werden.

4.3.1 Planabweichungen

Im Zusammenhang mit der Prioritätenvergabe sind wesentliche Planabweichungen entweder Differenzen bzgl. der Ankunfts- bzw. Abfahrtszeit eines Zuges oder das Nichteinhalten von vorgesehenen Anschlüssen gegenüber dem ursprünglichen Sollfahrplan. Aus der Gesamtheit der Planabweichungen in Kombination mit dem ursprünglichen Sollfahrplan läßt sich ein Dispositionsfahrplan ableiten, der den tagesaktuell gefahrenen Verkehr repräsentiert. Diese Planabweichungen werden später durch das Bonus-Malus-System bewertet und ggf. mit einer Bonus-Malus-Zahlung pönalisiert. Entsprechend ergibt sich das in Abb. 4-13 stehende, weitgehend selbsterklärende ER-Diagramm.

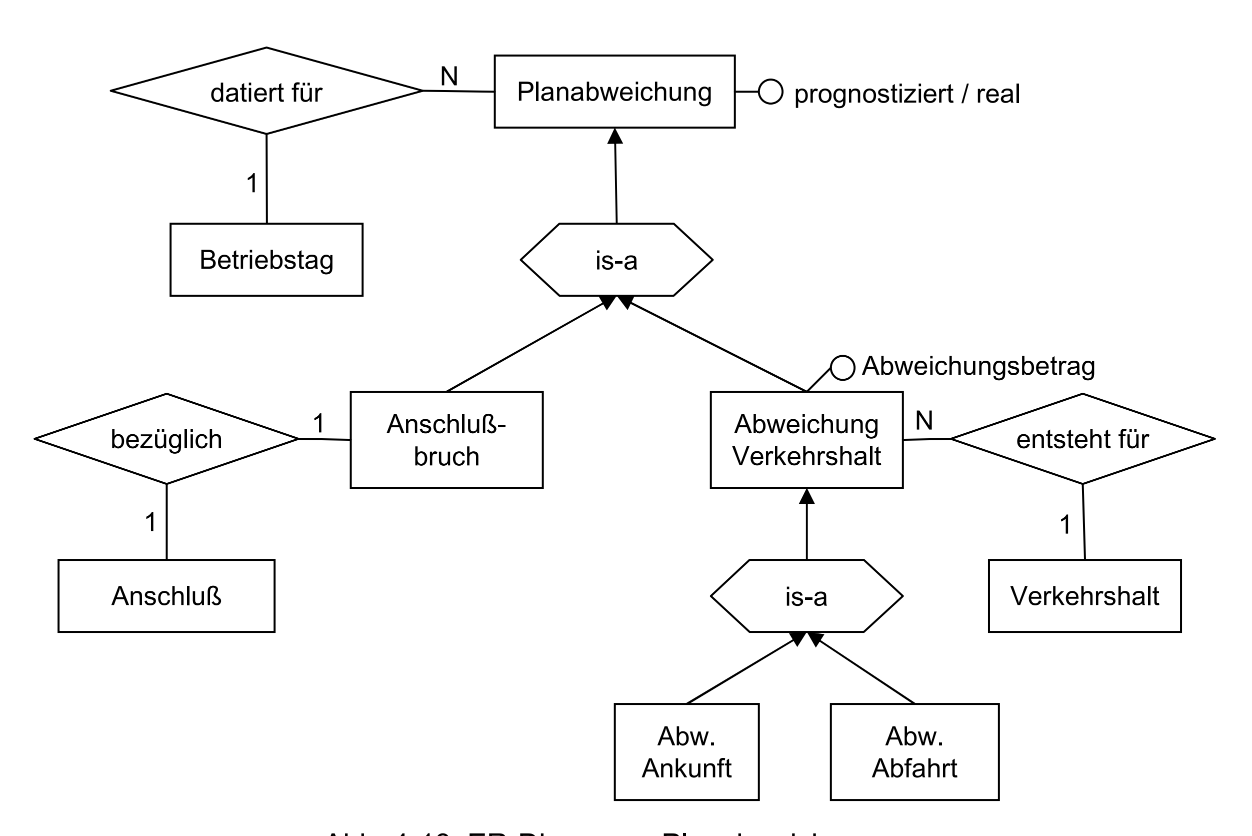


Abb. 4-13: ER-Diagramm Planabweichungen

Wichtig ist dabei, daß der Zustand „Planabweichung“ aufgrund von Prognoserechnungen und Auslösung eines Meldungsereignisses bereits schon vor seinem konkreten Eintreten bekannt sein kann und Prioritäten dann auf Basis dieser Prognose festgelegt werden können. Daher wird hier ein Attribut eingefügt, welches darüber Auskunft gibt, ob es sich beim korrespondierenden Meldungsereignis um ein prognostiziertes oder bereits konkret stattgefundenes Ereignis handelt. Falls das Ereignis in der Vergangenheit liegt, kann der Zustand verständlicherweise nicht mehr terminiert werden. Bei einem prognostizierten Ereignis sind die Zustandsattribute natürlich ebenfalls nur als prognostiziert zu betrachten. Der Zustand

kann bzgl. des Abweichungsbetrags immer wieder durch weitere Meldungsereignisse aktualisiert werden. In seltenen Fällen kann es auch sein, das die von einem Dispositionssystem prognostizierte Planabweichung doch nicht eintrifft, worauf eine Terminierung des Zustandes möglich ist.

4.3.2 Störungen

Störungen kann man als einen bestimmten Teilzustand des Eisenbahnsystems ansehen, infolgedessen (vom Typ der Störung abhängige) bestimmte Einschränkungen des Eisenbahnbetriebs resultieren, die unter Umständen geänderte Prioritäten erfordern. Wieder wird versucht, die Anzahl der verschiedenen Störungstypen minimal zu halten, um prioritätenbasierte Reaktionen beherrschbar zu halten. Man kann zwischen punktuellen Störungen und kontinuierlichen Störungen unterscheiden. Punktuelle Störungen betreffen einen bestimmten Zug, der an irgendeinem Punkt im Eisenbahnnetz aufgrund von technischen, organisatorischen oder externen (z. B. zu späte Beladung) Ursachen eine Verspätung erlitten hat und nach Abschluß dieser Störung ungehindert weiterfahren kann. Bei kontinuierlichen Störungen hingegen handelt es sich um eine Zug- oder eine Infrastrukturstörung, wobei hier ein längerfristiges technisches, den Betrieb beeinflussendes Problem im Vordergrund steht. Die Dauer dieser Störung ergibt sich damit aus dem Zeitintervall zwischen den korrespondierenden Initialisierungs- und Terminierungsereignissen.

Bereits bei Initialisierung der Störung wird eine Dauer der Störung prognostiziert. Falls eine Störung länger/kürzer andauert als prognostiziert, wird die Störung durch ein weiteres Störungsereignis aktualisiert. Eine kontinuierliche Zugstörung kann unabhängig vom eigentlichen Problem näherungsweise durch eine reduzierte Höchstgeschwindigkeit quantifiziert werden. Für Infrastrukturstörungen kann analog eine (geschätzte) prozentuale Kapazitätsreduktion des Streckenabschnitts bzw. des Infrastrukturknotens zur Quantifizierung verwendet werden. Für Streckenstörungen spielt außerdem eine Rolle, ob nur noch eingleisiger Betrieb möglich ist.

Ein weiterer Grund für die Unterteilung in Infrastruktur- und Zugstörungen sowie der Aufnahme des Verursacherattributes für punktuelle Störungen im Rahmen der Modellierung ist die sogenannte „Schuldfrage“ an aus einer Planabweichung möglicherweise resultierender Bonus-Malus-Zahlung. Läßt sich eine Planabweichung auf einen ursprünglichen Fehler des Infrastrukturmanagers rückverfolgen, dann ist dieser dafür „haftbar“, ansonsten gehen die Kosten auf den Zugbetreiber über. Sind mehrere Störungen in gleicher Weise ursächlich findet entsprechend eine Kostenteilung statt. Bei externen Störungen können Bonus-Malus-Regelungen komplett entfallen. Ein Verfahren zur Aufteilung von Bonus-Malus-Zahlungen auf verschiedene Verursacher wird in Kapitel 5 dargestellt. Dafür werden als Input Informationen über die angefallene Zusatzverspätung sowie die Gesamtverspätung des Zuges bei Abschluß der Störung benötigt, die entsprechend ebenso in das Modell aufgenommen werden sollen. Aus der gesamten Diskussion ergibt sich das ER-Schema in Abb. 4-14.

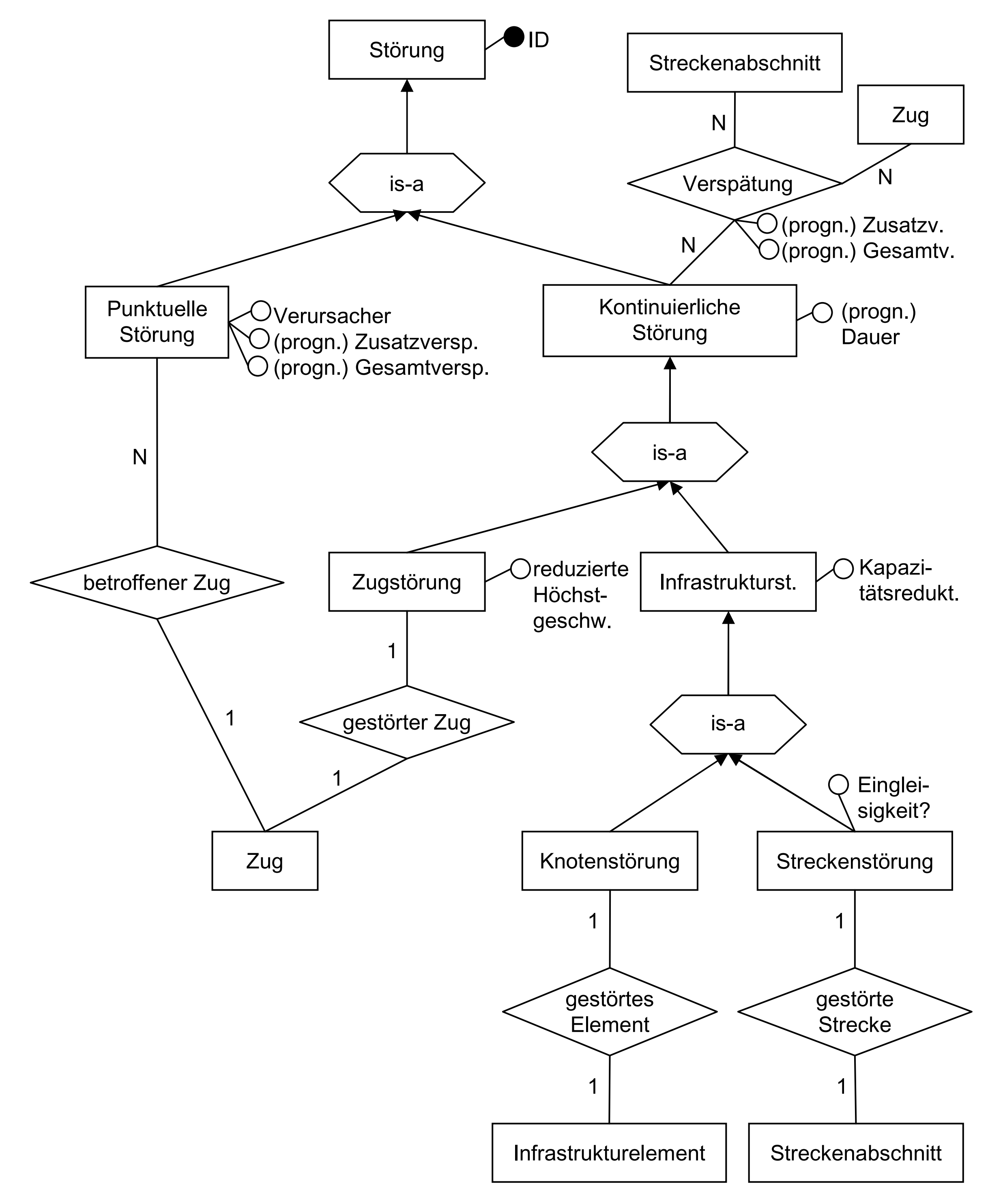


Abb. 4-14: ER-Diagramm Störungen

4.3.3 Prioritätenbasierte Dispositionsmaßnahmen

Ein weiterer wesentlicher Aspekt der Ereignissicht sind die zugspezifischen prioritätenbasierten Dispositionsmaßnahmen, die als Reaktionen auf konkrete Konflikte im Betrieb durchgeführt werden. Die Konflikte können in Belegungskonflikte (zwischen zwei Zügen auf einer Strecke) und Anschluß- bzw. Umlaufkonflikte (Übergänge zwischen zwei Zügen im Bahnhof) unterteilt werden. Bei diesen Dispositionsmaßnahmen handelt es sich um eine praktische

Umsetzung von Zug- und Anschlußprioritäten in konkrete Maßnahmen zur Behebung der Konflikte. Man unterscheidet hier zwischen Fahrzeitänderung, Haltezeitverlängerungen und Haltverlegungen. Bei Belegungskonflikten ist hier i. d. R. der Zug mit der niedrigeren Abschnittsweisen Zugpriorität betroffen. Ferner soll die Entscheidung, einen Anschluß nicht abzuwarten, ebenfalls als prioritätenbasierte Dispositionsmaßnahme modelliert werden. Im Falle von Umlaufverknüpfungen kann dies eine Verkürzung von Laufwegen zur Folge haben. Aus den Dispositionsmaßnahmen können auch Planabweichungen resultieren, die wiederum zu Bonus-Malus-Zahlungen führen können. Die Dispositionsmaßnahmen führen daher zu einem veränderten Dispositionsfahrplan, nach dem der Eisenbahnbetrieb dann konkret abgewickelt wird. Ziel des Dispositionsfahrplanes ist es, die Bonus-Malus-Zahlungen und damit die korrespondierenden Planabweichungen gegenüber dem Sollfahrplan möglichst gering zu halten.

Die prioritätenbasierten Dispositionsmaßnahmen sind ein spezielles Ereignis und werden zugspezifisch modelliert. Wiederum werden zur späteren Aufteilung von Bonus-Malus-Zahlungen Angaben zur Zusatz- und zur Gesamtverspätung durch eine prioritätenbasierte Dispositionsmaßnahme aufgenommen. Details können dem ER-Diagramm in Abb. 4-15 entnommen werden.

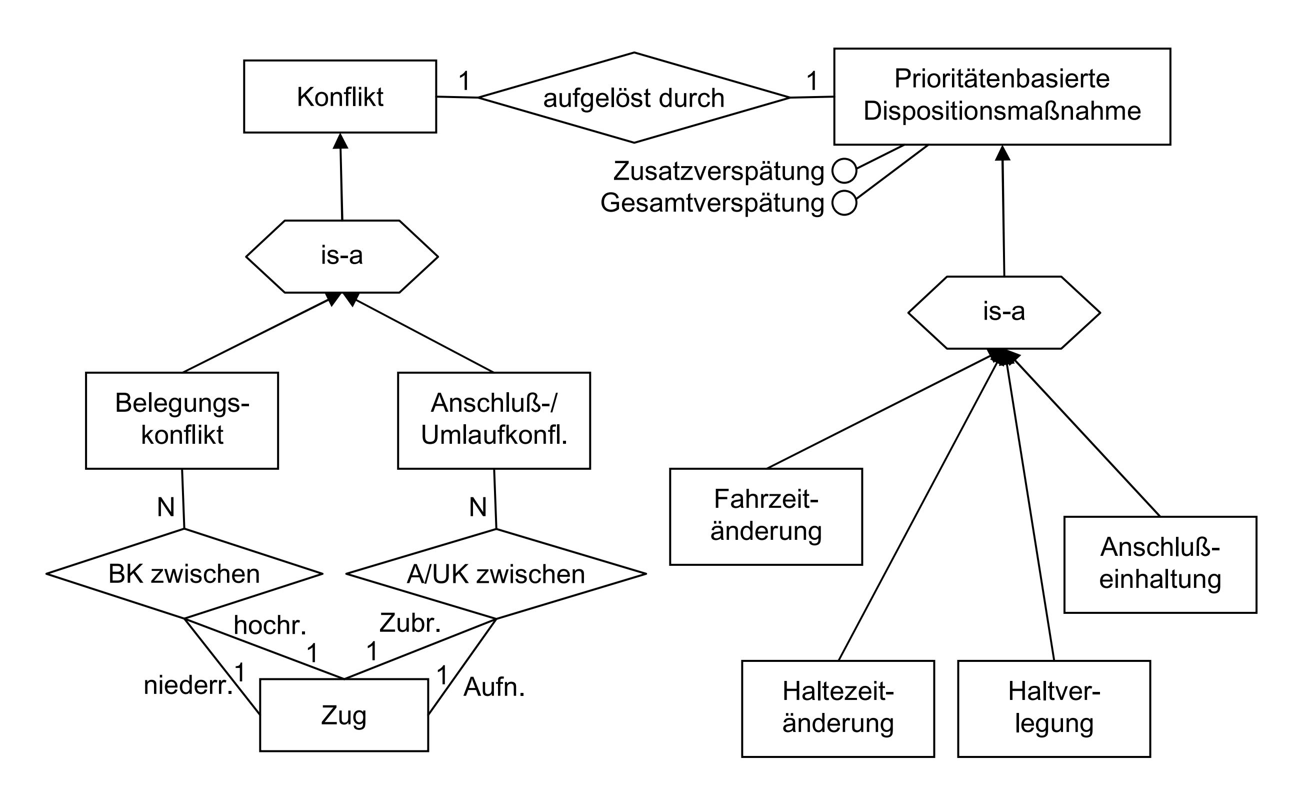


Abb. 4-15: ER-Diagramm Dispositionsmaßnahmen

4.3.4 Bonus-Malus-Zahlungen

Ein in dieser Arbeit betrachtetes Bonus-Malus-System dient dazu, im Falle von Planabweichungen (in der Regel Verspätungen bei bestimmten Verkehrshalten oder Brechung bestimmter Anschlüsse) Bonus-Malus-Zahlungen entsprechend realer verkehrswirtschaftlicher Bewertungen von Planabweichungen festzulegen. Dabei wird dem Abnehmer einer Ver-

kehrsleistung seitens des Anbieters der Leistung ein finanzieller Ausgleich für aus einer Planabweichung resultierende Qualitätseinbußen gewährt. Bei Verspätungen ist die Bewertungsfunktion abhängig von deren Betrag (i. d. R. in Minuten), bei Anschlußverlusten wird ein Pauschalbetrag festgelegt. Die Bonus-Malus-Bewertungsfunktionen wurden bereits in der Fahrplansicht bzgl. Verkehrshalten und Anschlüssen modelliert.

In dieser Arbeit wird eine Adäquatheit von Bonus-Malus-Systemen auf dem Primär- und Sekundärmarkt ausgegangen, d. h. die Erstattungen der Verkehrsunternehmen gegenüber den Endkunden für Planabweichungen werden intern ebenso zwischen dem Verkehrsunternehmen und dem Infrastrukturbetreiber entsprechend des „Verursacherprinzips“ verrechnet. Ein genaues Verfahren dazu wird in Kapitel 5 vorgestellt. Der Verfasser geht davon aus, daß sich dieses Prinzip (unabhängig von der aktuellen Situation zum Zeitpunkt der Veröffentlichung) langfristig in der Praxis durchsetzen wird.

Zu jeder Bonus-Malus-Zahlung aufgrund einer Bonus-Malus-Vereinbarung gibt es folglich eine zugehörige Planabweichung.

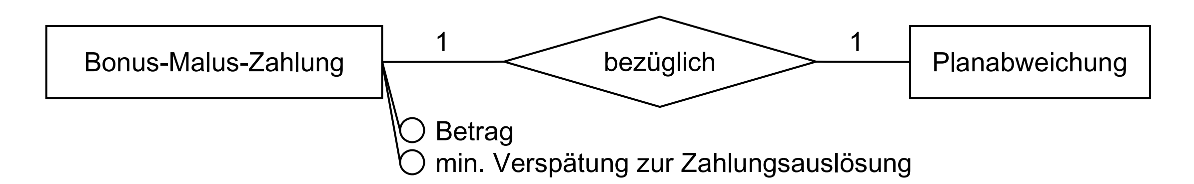


Abb. 4-16: ER-Diagramm Bonus-Malus-Zahlungen

Von besonderem Interesse ist neben dem Betrag der Bonus-Malus-Zahlung auch die minimale Verspätung des verspäteten Zuges bzw. Anschlußzubringerzuges, bei der die entsprechende Zahlung ausgelöst wird, um das Optimierungspotential durch Prioritäten entsprechend beurteilen zu können. Es ergibt sich das ER-Diagramm in Abb. 4-16.

Mit dem Ziel, derartige Bonus-Malus-Zahlungen zu minimieren, soll basierend auf der hier gezeigten Modellierung des Eisenbahnbetriebs im folgenden Kapitel ein Verfahren zur Prioritätenvergabe und -optimierung entwickelt werden.

5 Ein neues Verfahren zur Prioritätenvergabe und -optimierung

Im aktuellen Stand von Wissenschaft und Technik erfolgt eine prioritätenbasierte Disposition des Eisenbahnbetriebes in aller Regel aufgrund manuell vorgegebener statischer Prioritäten (vgl. 2.4). Damit läßt sich eine Minimierung von Bonus-Malus-Zahlungen, insbesondere bei komplexen verkehrswirtschaftlichen Zusammenhängen, nicht realisieren. Geschlossene Optimierungsverfahren hingegen lassen sich in den zur Verfügung stehenden Reaktionszeiten meist nicht realisieren. Daher soll im Rahmen dieses Kapitels nun ein neues und innovatives Verfahren zur Prioritätenvergabe und -optimierung entwickelt werden. Dieses Verfahren versucht die Qualität der prioritätenbasierter Dispositionsverfahren durch Verwendung kontinuierlich bzgl. dem dahinterliegendem Bonus-Malus-System angepaßter, dynamischer Prioritäten deutlich zu verbessern, wobei der Vorteil der schnellen Antwortzeiten prioritätenbasierter Verfahren erhalten bleibt. Die Prioritätenoptimierung wird dabei als „Hintergrundprozeß“ unabhängig vom Tagesgeschäft der Disposition durchgeführt. Die vorab in diesem Prozeß optimierten Zug- und Anschlußprioritäten können dem Disponenten dann bei Bedarf „ad hoc“ zur kurzfristigen Problemlösung mitgeteilt werden.

5.1 Grundlagen des vorgeschlagenen Verfahrens

Das hier vorgeschlagene Verfahren zur Prioritätenvergabe und -optimierung dient primär dem Ziel, die Gesamtsumme von aufgrund von Planabweichungen induzierten Bonus-Malus-Zahlungen zu minimieren (vgl. Ausführungen in 4.3.4). Daher müssen die Prioritäten in einer Weise festgesetzt werden, daß die darauf basierenden Dispositionsentscheidungen zu den entsprechend gewünschten Reaktionen führen.

5.1.1 Standard- und Störfallprioritäten

Im Interesse einer späteren Akzeptanz soll eine relative Monotonie der erlangten Lösungen angestrebt werden, d. h. man sollte von der Möglichkeit zu spontanen Prioritätenwechseln zurückhaltend Gebrauch machen. Insbesondere plötzliche Wechsel der Abschnittsweisen Zugprioritäten sollten möglichst vermieden werden oder auf gewichtige Störfälle beschränkt bleiben.

Daher sollte zwischen dem Fall eines weitgehend ungestörten Betriebs und bestimmten Störfallszenarien unterschieden werden. Wenn beispielsweise ein Zug einen technischen Defekt aufweist, sollte dies bei der Wahl der Prioritäten (durch Absenkung der Priorität für diesen Zug) berücksichtigt werden. Auch sollte ggf. auf eine größere Verspätung eines Zuges reagiert werden, um die Gesamtbilanz der Bonus-Malus-Zahlungen zu optimieren.

Dazu soll nun der Begriff des „Speziellen Betriebszustandes“ eingeführt werden. Bei einem Speziellen Betriebszustand wird entweder durch eine „kritische“ Planabweichung bei Ankunft im „Grenzbahnhof“ vor der Einfahrt in den Prioritätenabschnitt oder durch eine terminierte punktuelle bzw. andauernde kontinuierliche Störung verursacht. Konkret wird er durch Be-

triebsereignisse initialisiert, terminiert oder aktualisiert (siehe Abb. 5-1). Ähnliche Spezielle Betriebszustände können in „Speziellen Betriebszustandstypen“ gruppiert werden.

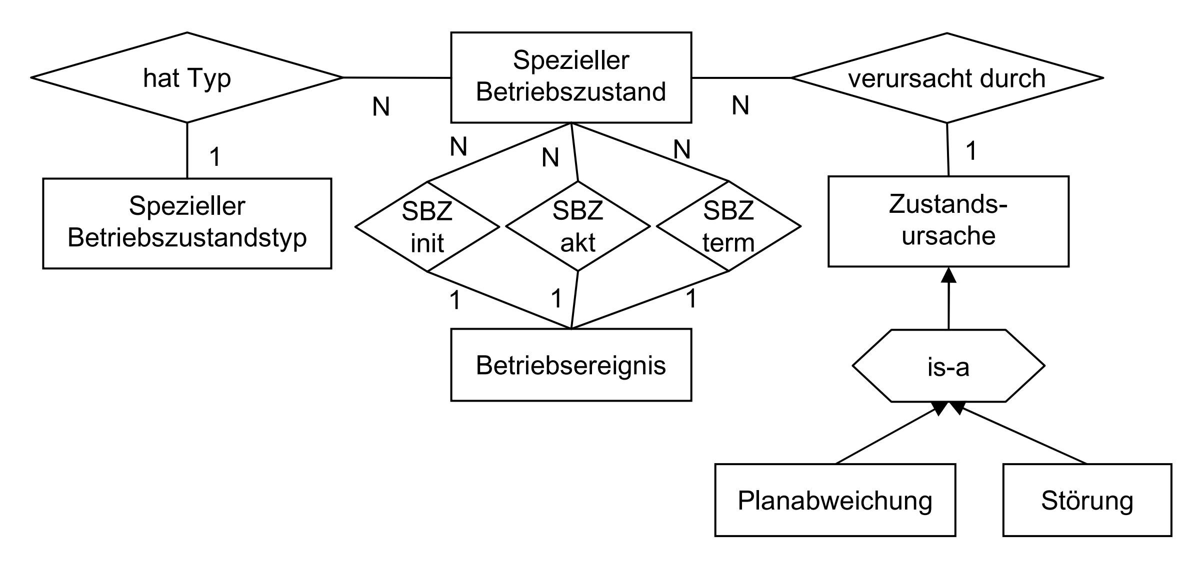


Abb. 5-1: ER-Diagramm Spezieller Betriebszustand

Man kann im Rahmen der Prioritätenoptimierung zwischen Standard- und Störfallprioritäten unterscheiden. Standardprioritäten sind die für jeden Zeitpunkt und Ort fest definierten, im Falle von ungestörtem bzw. nur leicht gestörtem Verkehr zu verwendenden Abschnittsweisen Zug- und Anschlußprioritäten. Bei Störfallprioritäten handelt es sich um für bestimmte vorab definierte Typen von Speziellen Betriebszuständen vorgegebene Abweichungen von den genannten Standardprioritäten („Notfallpläne“). Dem Disponenten wird im hier vorgestellten Verfahren grundsätzlich die jeweils gerade zutreffende Standard- oder Störfallpriorität zugewiesen. Eine Störfallpriorität kann ggf. gleichzeitig mit mehreren Speziellen Betriebszustandstypen mit gleicher Semantik assoziiert sein.

Die entsprechend erweiterten ER-Modelle für Abschnittsweise Zug- und Anschlußprioritäten finden sich in Abb. 5-2 und 5-3.

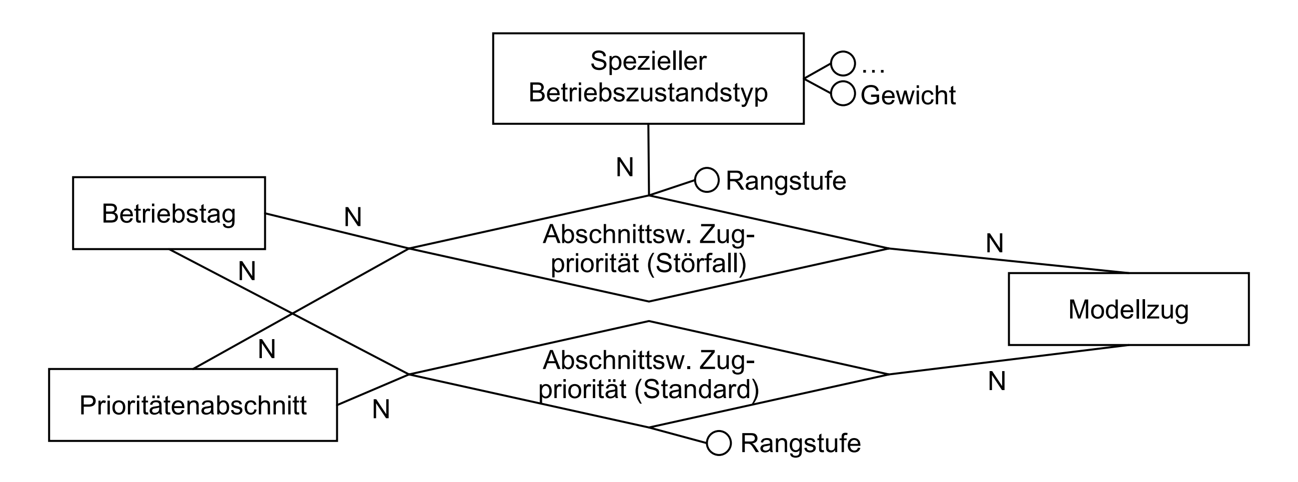


Abb. 5-2: ER-Diagramm Standard- und Störfallprioritäten für Abschnittsweise Zugprioritäten

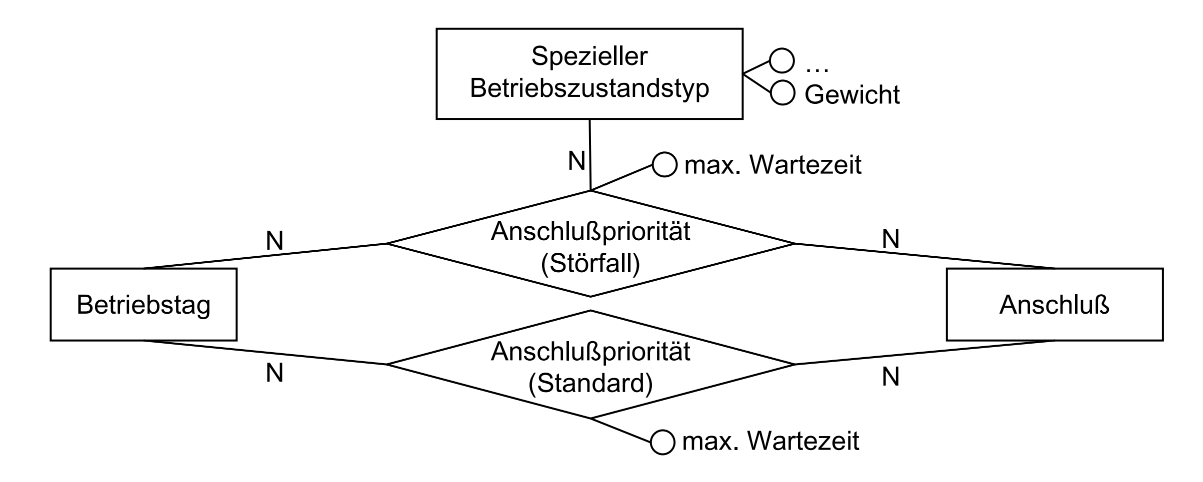


Abb. 5-3: ER-Diagramm Standard- und Störfallprioritäten für Anschlußprioritäten

Mit der Unterscheidung zwischen Standard- und Störfallprioritäten und der entsprechenden Vergabe der passenden, vorab für die entsprechende Konstellation ermittelten Priorität ist bereits eine Form von Optimierung entsprechend einer gegebenen Betriebssituation vorgegeben. Dazu muß festgelegt werden, wie viele verschiedene im Zusammenhang mit Störfallprioritäten relevanten Speziellen Betriebszustandstypen vorgegeben werden.

Im Sinne einer späteren Akzeptanz aller am Betrieb Beteiligten sollen die Störfallprioritäten im Zusammenhang mit Planabweichungen eine „Neutralität" gegenüber dem eigentlichen Verursacher derselben aufweisen. Auf 20 Minuten Verspätung aufgrund einer vom Infrastrukturbetreiber zu vertretenen Signalstörung soll in gleicher Weise, wie auf 20 Minuten Verspätung aufgrund vom Verkehrsunternehmen zu vertretener Rangierprobleme reagiert werden. Entsprechend sollten hier generische Spezielle Betriebszustandstypen berücksichtigt werden, für die dann jeweils eine einheitliche Störfallpriorität vorgehalten wird. Die Anzahl der verschiedenen Speziellen Betriebszustandstypen bzw. Störfallprioritäten soll möglichst gering gehalten werden und im Zusammenhang mit den jeweiligen Abstufungen im Bonus-Malus-System und der Möglichkeit zur Verspätungsreduktion stehen.

Dies soll im folgenden genauer erläutert werden: Für einen Modellzug MZ sind mehrere Bonus-Malus-Regelungen BM_1, BM_2, …, BM_n für Planabweichungen an einem oder mehreren Verkehrshalten definiert. Hierbei handelt es sich um Zahlungen, die an Verkehrshalten ab einem bestimmten Verspätungsbetrag bzw. einem Anschlußbruch anfallen. Jeder durch Planabweichungen verursachte Spezielle Betriebszustandstyp steht in Relation zur Gefährdung einer möglichen Bonus-Malus-Regelung BM_i. SBZ(MZ,PA) entspricht die Menge der für MZ auf dem Prioritätenabschnitt PA relevanten Speziellen Betriebszustandstypen. Allgemein kann man für die Festlegung der Verspätungsintervalle $[x_i,y_i]$ eines die Gefährdung von BM_i implizierenden, durch eine Planabweichung verursachten Speziellen Betriebszustandstyps $SBZ_i \in SBZ(MZ,PA)$ auf einem Prioritätenabschnitt PA sagen, daß y_i dem Verspätungsbetrag von MZ bei Ankunft im „Grenzbahnhof" vor der Einfahrt in PA entspricht, bei dem die Verletzung von BM_i durch Dispositionsmaßnahmen für Belegungskonflikte nicht mehr verhindert werden kann. Der Wert von x_i ist etwas schwieriger allgemeingültig festzu-

legen. Dieser Wert spiegelt die unmittelbare Gefahr einer Verletzung von BM_i im Falle weiterer für einen Zug ungünstiger Dispositionsmaßnahmen wieder. Dabei spielen die Länge des Prioritätenabschnitts und damit die Wahrscheinlichkeit für Belegungskonflikte in diesem Prioritätenabschnitt eine Rolle. Dabei kann auf historisches Datenmaterial zurückgegriffen werden. Die Länge des Intervalls $[x_i,y_i]$ soll proportional zum Erwartungswert für Belegungskonflikte $E_{BK}(MZ,PA)$ gewählt werden. Pro erwarteten Belegungskonflikt sollen die mittlere durch eine prioritätenbasierte Dispositionsmaßnahme verursachte Folgeverspätung $E_{FV}(MZ,PA)$ berücksichtigt werden. Damit ergibt sich

$$x_i = y_i - E_{BK}(MZ,PA) \cdot E_{FV}(MZ,PA)$$

Ferner soll es für den Modellzug MZ auf einem Prioritätenabschnitt PA einen die unvermeidliche Auslösung aller Bonus-Malus-Zahlungen implizierenden, durch eine Planabweichung verursachten Speziellen Betriebszustandtyp mit dem Intervall $[x_z, \infty]$ geben, wobei x_z dem Verspätungszeitpunkt entspricht, bei dem das Maximum an Bonus-Malus-Zahlungen für MZ ausgelöst wird und auf keine Weise mehr durch Dispositionsmaßnahmen vermeidbar ist. Wenn z. B. alle Verspätungen größer als 30 Minuten im Bonus-Malus-System mit dem gleichen Betrag abgerechnet werden, macht es keinen Sinn, zwischen dreistündigen und vierstündigen Verspätungen bzgl. Störfallprioritäten zu differenzieren, wenn die Zahlung ohnehin nicht mehr verhindert werden kann. Störfallprioritäten im Zusammenhang mit Planabweichung sind primär für Abschnittsweise Zugprioritäten von Bedeutung, da (zulässige) Verspätungsbeträge in Anschlußprioritäten bereits explizit berücksichtigt werden und somit keine weiteren Spezialisierungen mehr benötigt werden.

Sinngemäß können auf gleiche Weise auch durch nach der Einfahrt in den Prioritätenabschnitt auftretende terminierte punktuelle Störungen verursachte Spezielle Betriebszustandstypen festgelegt werden. Die wiederum die Gefährdung einer möglichen Bonus-Malus-Regelung BM_i implizierenden Speziellen Betriebszustandstypen für punktuellen Störungen können auf Basis einer minimalen unmittelbar durch die Störung bewirkten Zusatzverspätung und wiederum eines Intervalls einer am Ende der Störung resultierten Gesamtverspätung definiert werden. Auf eine exakte Formalisierung wird an dieser Stelle wegen der Ähnlichkeit zu Planabweichungen verzichtet.

Man beachte, daß auf diese Weise zwei Spezielle Betriebszustandstypen die Gefährdung von BM_i implizieren können, einmal verursacht durch Planabweichungen und einmal verursacht durch punktuelle Störungen. In diesem Sinne soll hierfür auch nur eine Störfallpriorität definiert werden, die gleichzeitig mit beiden Speziellen Betriebszustandstypen verknüpft ist.

Dieser „neutrale“ Ansatz kann im Zusammenhang mit Störfallprioritäten für andauernde kontinuierliche Störungen so nicht durchgeführt werden. In diesem Falle müssen individuelle Spezielle Betriebszustandstypen und Störfallprioritäten für die gemäß 4.3.2 modellierten kontinuierlichen Störungen spezifiziert werden. Um das Verfahren langfristig wartbar zu halten, soll allerdings die Anzahl der Zustandstypen möglichst gering gehalten werden, z. B. „Reduzierte Zuggeschwindigkeit zwischen x und y km/h“, „Reduzierte Streckengeschwindigkeit zwischen x und y km/h“, „Eingleisiger Betrieb“ etc. Die Anzahl der benötigten Zustandstypen und deren Geschwindigkeitsintervalle für reduzierte Zug- und Streckengeschwindig-

keiten hängt insbesondere mit den Geschwindigkeitsprofilen der verschiedenen auf einem Prioritätenabschnitt fahrenden Modellzüge zusammen. Bei starker Streuung sind mehr Typen erforderlich als bei einem weitgehend bzgl. Geschwindigkeit einheitlichen Betrieb. Falls ein relativ schneller Zug technisch nicht mehr in der Lage ist, wirklich schneller als ein relativ langsamerer Konkurrent zu fahren, sollte dies in einer besonderen Störfallpriorität für den Zug abgebildet werden. Diese Arten von Störfallprioritäten sind sowohl für Abschnittsweise Zugprioritäten wie auch für Anschlußprioritäten von Bedeutung. Bei Anschlußprioritäten betreffen die Speziellen Betriebszustandstypen den anschlußaufnehmenden Zug bzw. von diesem befahrene Streckenabschnitte.

Die Speziellen Betriebszustandstypen haben ein unterschiedliches Gewicht, um im Fall von sich überlappenden Speziellen Betriebszuständen (z. B. ein defekter Zug auf einer Langsamfahrstelle) den maßgeblichen Betriebszustandstyp und damit eindeutig die maßgebliche Störfallpriorität bestimmen zu können. Die Gewichtung verschiedener Spezieller Betriebszustande bzgl. der Gefährdung von Bonus-Malus-Vereinbarungen wegen Verspätungen bzw. punktueller Störungen erfolgt auf Basis einer Betrachtung der jeweils assoziierten Bonus-Malus-Regelungen BM_i. Hierzu wird die daraus resultierende Bonus-Malus-Steigerung BMS_i betrachtet. Bei der Regelung BM_1 bzgl. eines Anschlußbruchs (Verspätung des Zubringers) entspricht der Betrag vom BM_1 auch BMS_1. Gleiches gilt wenn für eine Planabweichung nur eine Regelung BM_1 ab einer Verspätung von x Minuten vereinbart wurde. Bei Vorhandensein einer weiteren Regelung BM_2 ab einer Verspätung von y Minuten (mit $y>x$) errechnet sich BMS_2 als Differenz der Zahlungen für BM_2 und BM_1. Das Gewicht eines Speziellen Betriebszustandstyps steht in Relation zum Wert von BMS_i des assoziierten BM_i. Spezielle Betriebszustände bzgl. kontinuierlicher Störungen sind generell als gewichtiger gegenüber solchen wegen Planabweichungen bzw. punktueller Störungen einzustufen. Ferner soll im Sinne der Eindeutigkeit des maßgeblichen Speziellen Betriebszustandes der Typ des zeitlich zuerst eintreffenden Spezielle Betriebszustandes bzgl. der Planabweichung ein leicht höheres Gewicht als der bzgl. der Störfallpriorität „äquivalente“ Spezielle Betriebszustandstyp bzgl. der punktuellen Störung. Innerhalb der Speziellen Betriebszustandstypen bzgl. kontinuierlicher Störungen muß eine individuelle Ordnung erfolgen.

5.1.2 Prioritätenvergabe und -optimierung

Bei der Entscheidung, ob für einen Zug eine Standard- oder Störfallpriorität verwendet wird, handelt es sich um eine Ad-hoc-Optimierung im Rahmen der Disposition. Dieser Prozeß wird im folgenden auch als Prioritätenvergabe bezeichnet, wobei hier noch einmal zwischen einer planmäßigen und der störfallbedingten Prioritätenvergabe unterschieden wird.

Eine **planmäßige Prioritätenvergabe** bezieht sich auf einem vorab geplanten ortsgebundenen Wechsel der Zugpriorität im Rahmen der Disposition beim Übergang zwischen zwei Prioritätenabschnitten. Dabei spielt es keine Rolle, ob es sich bei der vergebenen Priorität um Standard- oder Störfallprioritäten handelt. Wenn ein Zug bei einem Übergang zwischen zwei Prioritätenabschnitten keinen Planhalt aufweist, kann bedingt durch eine planmäßige Prioritätenvergabe ggf. ein außerplanmäßiger Halt entstehen. Bei Vorhandensein eines

Planhalts kann aus der planmäßigen Prioritätenvergabe ggf. eine Haltezeitverlängerung resultieren.

Eine **störfallbedingte Prioritätenvergabe** hingegen entspricht kurzfristig auftretenden, temporären, (störfall-)ereignisbedingten Veränderungen der bei der Disposition verwendeten Abschnittsweisen Zug- und Anschlußprioritäten zwischen Standard- und Störfallprioritäten bzw. zwischen verschiedenen Störfallprioritäten (Neubewertung der betrieblichen Lage, z. B. wegen punktueller Störung oder aufgrund eingeschränkter Ressourcen). Die Veränderung der Prioritäten bei einem Eintritt bzw. beim Verlassen eines Speziellen Betriebszustands eines Typs mit hohem Gewicht wird bereits im Vorfeld festgelegt und gespeichert. Dadurch wird sichergestellt, daß keine umfangreichen und zeitkritischen Ad-hoc-Berechnungen bzw. Optimierungen „unter rollendem Rad“ durchgeführt werden müssen. Falls sich mehrere Spezielle Betriebszustände überlappen, ist der Typ mit höchstem Gewicht maßgeblich, so daß eine eindeutige Störfallpriorität festgelegt werden kann. Störfallbedingte Prioritätenvergaben sollen grundsätzlich nur bei schweren Störungen (hohes Gewicht des Speziellen Betriebszustandstyps) durchgeführt werden (hohe Bonus-Malus-Regelung gefährdende punktuelle Störungen oder stark eingeschränkte Ressourcen).

Neben der Prioritätenvergabe existiert aber noch eine andere Art der Optimierung, nämlich die längerfristige Ermittlung der jeweils am besten bereitzustellenden Standard- und Störfallprioritäten. Diese soll aufgrund protokollierter Bonus-Malus-Zahlungen tatsächlich gefahrener Dispositionsfahrpläne erfolgen. Während es sich bei der oben beschriebenen Prioritätenvergabe hauptsächlich um ein Auslesen und Übermitteln von Daten (Ereignisse/Betriebszustände und vorab festgelegte mögliche Prioritäten) handelt, ist der Prozeß der Prioritätenoptimierung der Kern des im Rahmen dieser Arbeit entwickelten Prioritätenvergabe- und -optimierungsverfahrens (vgl. Abb. 5-4).

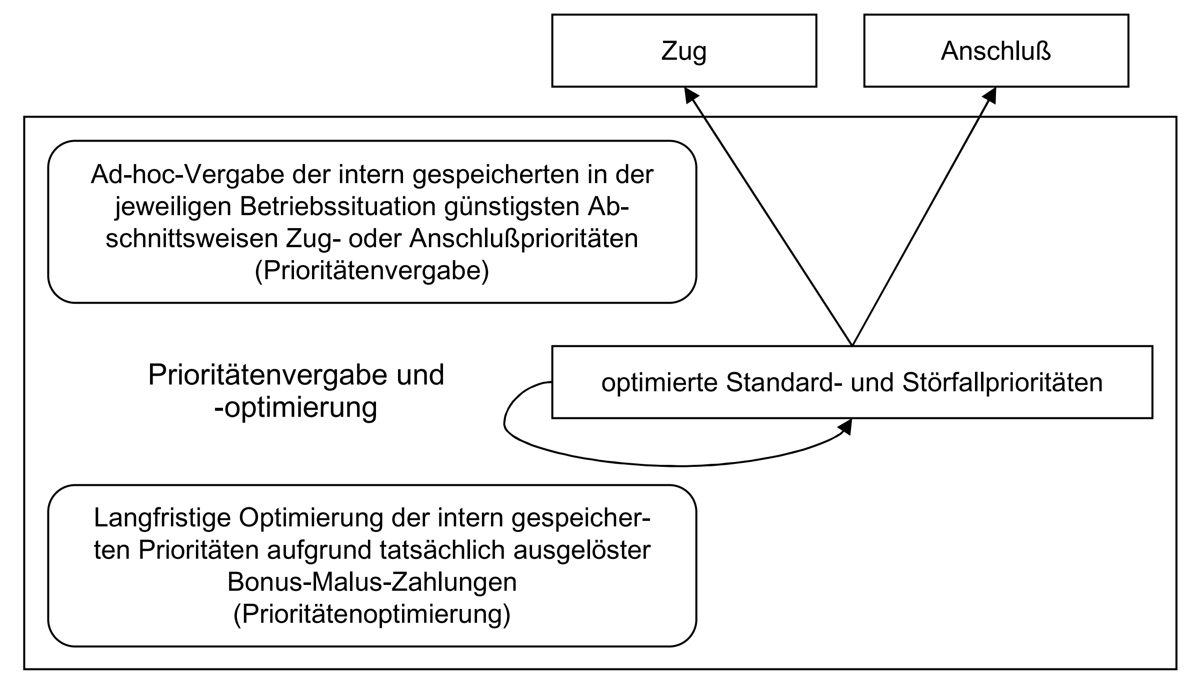


Abb. 5-4: Arten von Optimierungen (schematische Darstellung)

Bei der Prioritätenoptimierung wird auf konkret resultierende Bonus-Malus-Zahlungen im anhand der vergebenen Prioritäten disponierten Eisenbahnbetrieb reagiert. Wenn hier bestimmte Toleranzwerte über einen bestimmten Zeitraum kontinuierlich (Forderung nach relativer Monotonie) überschritten werden, wird der Prioritätenoptimierungsprozeß aktiviert. Anschließend wird das System versuchen, die Resultate durch dauerhafte Änderung der zukünftig bereitgestellten Standard- und/oder Störfallprioritäten entsprechend zu verbessern. Dies kann durch eine höhere Priorität von durch Bonus-Malus-Zahlungen besonders negativ betroffenen Modellzügen bzw. Anschlüssen oder durch eine Abwertung von in der Vergangenheit sehr unkritischen, aber andere Modellzüge und Anschlüsse behindernden Modellzügen und Anschlüssen erfolgen. Die genauen Details des Verfahrens sollen an dieser Stelle zunächst noch offen bleiben.

Bei dieser Prioritätenoptimierung müssen gesetzliche oder sonstige normative Vorgaben berücksichtigt werden. So kann es sein, daß gemäß dem Bonus-Malus-System ökonomisch optimale Prioritäten im Sinne der normativen Vorgaben nicht zulässig sind. Daher können für Abschnittsweise Zug- und Anschlußprioritäten zulässige Wertebereiche (ggf. differenziert in Standard- und Störfallprioritäten) vordefiniert werden, die dann im Rahmen der Prioritätenoptimierung in jedem Falle eingehalten werden müssen. Auf diese Weise können z. B. gesetzlich vorgeschriebene Vorrangstrecken für bestimmte Verkehre im Verfahren erzwungen werden.

Die Grundphilosophie der langfristigen Prioritätenoptimierung ist somit, daß sich die vom System in einem entsprechenden Verfahren optimierten Standard- und Störfallprioritäten für Modellzüge und Anschlüsse anschließend bei zukünftigen Dispositionsentscheidungen im realen Betrieb bewähren müssen. Aus zukünftigen Dispositionsergebnissen kann die Güte eines kartesischen Produktes verschiedener Prioritäten (Prioritätenvektor) bzgl. des vorgegebenen Bonus-Malus-Systems abgelesen werden. Wenn mittels einer nachträglichen Analyse des realen, protokollierten Verkehrs nachhaltige Probleme im Betrieb erkannt werden, muß versucht werden, in geeigneter Weise darauf durch maßvolle Prioritätenanpassungen zu reagieren. Im Gegensatz zur Prioritätenvergabe erfolgt die Prioritätenoptimierung nicht in Echtzeit und ist somit weitgehend zeitunkritisch. Es muß sichergestellt sein, daß die Resultate mittelfristig zu Gesamtverbesserungen führen.

Bei der Initialisierung des Verfahrens ohne stabiles Datenmaterial ist es unter Umständen ratsam, eine erste Eichung der Startprioritäten vor dem wirklichen Dispositionseinsatz vorzunehmen. Dazu können die Prioritäten zunächst einmal mit Hilfe eines Eisenbahnbetriebssimulators mit „zufällig" erzeugten Störungen (bei prioritätenbasierter Konfliktlösung) geprüft und stabilisiert werden. Im allgemeinen ist davon auszugehen, daß die meisten Prioritätenoptimierungen in der Anfangsphase (nach der Initialisierung der Prioritäten) auftreten. Wenn diese Anfangsphase bis zum Erreichen weitestgehend stabiler Prioritäten nicht im laufenden Betrieb, sondern bereits in einer Simulation erfolgt, bleibt der Betrieb von entsprechend „schwankenden" Dispositionsergebnissen verschont und man kann sich dort auf die wirkliche Feinjustierung im praktischen Einsatz unter realen Bedingungen beschränken.

5.2 Bewertung und Anpassung von Prioritäten im Rahmen der Optimierung

Eine Veränderung von Prioritäten im Rahmen der Prioritätenoptimierung kann einerseits zu Kosten, andererseits aber auch zu Nutzen führen, die gegeneinander abgewogen werden müssen. Eine lokale bzw. globale Prioritätenerhöhung eines Modellzuges wegen ungünstiger Bonus-Malus-Werte bzgl. desselben hat als Nutzen ggf. eine Reduktion dieser Bewertung zur Folge, kann aber als Kosten auch ungünstigere Bonus-Malus-Werte für andere (bzgl. Priorität relativ abgesenkte) Modellzüge implizieren. Gleiches gilt analog auch bei einer lokalen bzw. globalen Prioritätenabsenkung eines Modellzuges. Erhöhungen von Anschlußprioritäten wirken positiv auf die Anschlußsicherheit, aber negativ auf das Pünktlichkeitsniveau.

Als Qualitätsmaßstab der Disposition wird das zugrundeliegende Bonus-Malus-System verwendet. Ziel der Prioritätenoptimierung ist es, das bei auf dessen Ergebnissen resultierender Betriebsdisposition Bonus-Malus-Zahlungen vermieden oder zumindest soweit möglich reduziert werden. Im Rahmen der Disposition können die Prioritäten so allerdings nicht in jedem Falle 1:1 umgesetzt werden, da im Zweifelsfall Aspekte wie die Flüssigkeit des Verkehrs Vorrang vor durch das Bonus-Malus-System ausgedrückten kommerziellen Aspekten haben muß.

5.2.1 Verursachung von Bonus-Malus-Zahlungen

Es besteht die Möglichkeit, daß Bonus-Malus-Zahlungen

- aufgrund von betrieblichen Störungen am Fahrzeug oder Fahrweg bzw. externe Störungen oder
- aufgrund von prioritätenbasierenden Dispositionsmaßnahmen

entstanden sind.

Die erste Gruppe von Ereignissen kann durch prioritätenbasierte Dispositionsmaßnahmen nicht vermieden, sondern lediglich in ihren weiteren Folgen beschränkt werden, was im hier dargestellten Verfahren durch die Definition und Verwendung von Störfallprioritäten ermöglicht wird. Von besonderem Interesse für die Optimierung der jeweiligen Standard- und Störfallprioritäten sind daher jene Anteile, die unmittelbar aufgrund von prioritätenbasierten Dispositionsmaßnahmen (möglicherweise unnötig) entstanden sind und somit durch die prioritätenbasierte Disposition direkt beeinflußbar sind.

Da ein exaktes Verfahren zur Bestimmung von Ursachen von Bonus-Malus-Zahlungen sich als unverhältnismäßig rechenaufwendig darstellt, soll hier eine zufriedenstellende induktive Lösung favorisiert werden, die kurzfristig hinreichend genaue Schätzwerte bestimmen kann. Hierbei wird der Bereich zwischen dem letzten auf dem Laufweg vor einem Bonus-Malus-Meßpunkt gelegenen Punkt, an dem der Zug noch pünktlich verkehrte (bzw. der Startpunkt, falls eine durchgängige Verspätung vorlag), bis ggf. zum Bonus-Malus-Meßpunkt genauer betrachtet.

Bei der erstmaligen Verursachung einer Planabweichung in Form einer Verspätung eines Zuges auf seinem Laufweg wird die Verursachung der Bonus-Malus-Zahlung aufgrund einer Verspätung zu 100% dieser verursachenden Störung oder prioritätenbasierten Dispositionsmaßnahme zugeschrieben. Wenn sich, bedingt durch eine Verspätungsursache, eine bereits vorher vorhandene Verspätung eines Zuges erhöht, wird der Anteil der zusätzlich angefallenen Verspätung an der Gesamtverspätung als Schulanteil dieser Verspätungsursache an der Bonus-Malus-Zahlung betrachtet. Die Gesamtschuldanteile aller bereits an der Bonus-Malus-Zahlung zeitlich vorher eingetretenen mitschuldigen Störungen und Dispositionsmaßnahmen wird entsprechend auf den Anteil der bereits vorher bestehenden Verspätung an der neuen Gesamtverspätung gekürzt. Das interne Verursachungsverhältnis der bereits vorher bestehenden „Schuldigen“ bleibt durch die neu eintretende Verspätungsursache unberührt. Bzgl. Anschlußaufnahmen ist zu beachten, daß falls der Zug in einer Verkehrsstation als Aufnehmer für mehrere Anschlüsse dient, die Zuweisung des Bonus-Malus-Zahlungsanteil vollständig zu Lasten des eingehaltenen Anschlusses (prioritätenbasierte Dispositionsmaßnahme) erfolgt, welcher die längste Wartezeit des Aufnehmers verursacht hat.

Formal kann das im Rahmen der hier erstellten Modellierung folgendermaßen realisiert werden: Vor einer eine Bonus-Malus-Zahlung BM_x auslösende Planabweichung kann auf dem Zuglaufweg eine Sequenz mehrerer Verspätungsursachen $VU(BM_x) = [VU_1, VU_2, \ldots, VU_n]$ auftreten. Für alle diese VU_y, die im Detail Störungen oder prioritätenbasierte Dispositionsmaßnahmen sein können, ist gemäß der in Kapitel 4 dargestellten Modellierung die Zusatzverspätung $zv(VU_y)$ sowie die Gesamtverspätung am Ende der Ursache $gv(VU_y)$ bekannt. Für kontinuierliche Zugverspätungen wird im übrigen jeder Streckenabschnitt, auf dem die Störung andauert, als separater VU_y betrachtet. Ferner wird wie bereits gesagt bei einer mehreren parallelen Anschlußaufnahmen ausschließlich derjenige mit längster Wartezeit berücksichtigt. Es gilt entsprechend

$$zv(VU_1) = gv(VU_1) \quad \text{und} \quad zv(VU_i) < gv(VU_i),\ i > 1$$

Entsprechend gilt für den Anteil an einer Bonus-Malus-Zahlung wegen Verspätung $a_vs(BM_x, VU_y)$:

$$a_vs(BM_x, VU_y) = (1 - \sum_{VU_i \in VU(BM_x), i>y} a_vs(BM_x, VU_i)) \frac{zv(VU_y)}{gv(VU_y)}$$

Für Bonus-Malus-Zahlungen aufgrund gebrochener Anschlüsse ist einerseits die Ankunftsverspätung des Zubringerzuges von Bedeutung, die gemäß dem oben erwähnten Verfahren auf Verursacher umgelegt werden kann. Darüber hinaus muß noch die prioritätenbasierte Dispositionsmaßnahme, den Anschluß aufgrund einer zu geringen Anschlußpriorität nicht einzuhalten, berücksichtigt werden. Hierbei gilt bzgl. der Gewichtung der Anschlußpriorität (Wartezeit des aufnehmenden Zuges) gegenüber der Verspätung des Zubringerzuges das Prinzip, daß bei einer kleinen absoluten Ankunftsverspätung des Zubringerzuges am Anschlußbahnhof der Anteil der Anschlußpriorität höher gewichtet werden soll als bei einer hohen absoluten Verspätung des Zuges. Bei einer geringeren Verspätung liegt das Problem verstärkter an der zu niedrigen Anschlußpriorität, wohingegen bei einer hohen Verspätung

die Möglichkeit, die Anschlußqualität durch Verlängerung der Wartezeit des aufnehmenden Zuges nur mit großem Aufwand und hohen Kosten sichergestellt werden kann. Ebenso soll bei einer relativ großen Überschreitung der vorgegebenen Wartezeit die Anschlußpriorität relativ gering bewertet werden, als wenn der Anschluß nur um wenige Minuten verpaßt wurde.

Entsprechend sollen die Anteile der Verursachung der gesamten Bonus-Malus-Zahlung BM_X für einen Anschlußverlust $ANV \in VU(BM_x)$ (mit einer Anschluß-Wartezeit von $wz_{max}(AN)$ und einer Ankunftsverspätung vs_{ank}) folgendermaßen zwischen der Anschlußaufgabe a_avl(ANV) und den Verspätungsursachen des Zubringers $VU_y \in VU(BM_x) \setminus ANV$ aufgeteilt werden:

$$a_avl(BM_x, ANV) = \frac{wz_{max}(AN)}{vs_{ank}}$$

$$a_avl(BM_x, VU_y) = (1 - a_avl(BM_x, ANV))(1 - \sum_{VU_i \in VU(BM_x) \setminus ANV, i>y} a_avl(BM_x, VU_i))\frac{zv(VU_y)}{gv(VU_y)}$$

Das Verfahren zur Bestimmung der Verursacher einer Bonus-Malus-Zahlung soll bereits nach Durchfahrt durch einen Prioritätenabschnitt vorzeitig abgebrochen werden, sobald der gesamte auf dem Laufweg bisher angefallene Gesamtverspätungsbetrag den Mindestverspätungsbetrag $v_{min}(BM)$ zur Auslösung einer verspätungsbedingten Bonus-Malus-Zahlung durchgängig überschreitet bzw. der gebrochene Anschluß selbst bei Abwarten der vorgegebenen maximalen Wartezeit durchgängig nicht mehr eingehalten werden kann. Entsprechend wird die Sequenz $[VU_1, VU_2, \ldots, VU_n]$ auf eine Sequenz $VU(BM_x) = [VU_1, .., VU_i]$ verkürzt, wobei gilt:

$$\forall j, i \le j \le n : gv(VU_j) \ge v_{min}(BM) \text{ und } gv(VU_{i-1}) < v_{min}(BM)$$

Wenn die tatsächliche Bonus-Malus-Zahlung ab einer Verspätung von x Minuten ausgelöst wird, endet die Feststellung der Verursachungsanteile ab dem Zeitpunkt, zu dem der Zug durchgängig mit mindestens x Minuten Verspätung verkehrt. Noch weiter angesammelte Verspätung, die in keiner weiter erhöhten Bonus-Malus-Zahlung resultiert, soll nicht mehr bei der Anteilsermittlung für die Ursachen der Bonus-Malus-Zahlung berücksichtigt werden, da das eigentliche Problem auf dem bereits zu diesem Zeitpunkt passierten Laufwegsabschnitten liegt, und es ggf. aus Disponentensicht Sinn macht, einen stark verspäteten Zug entsprechend mit niedriger Priorität zu behandeln, wenn die Bonus-Malus-Zahlung ohnehin nicht mehr zu verhindern ist. Entsprechend wird das Aufteilungsverfahren lediglich für die Sequenz $VU(BM_x) = [VU_1, .., VU_i]$ angewendet.

Auf diese Weise lassen sich zu einer Bonus-Malus-Zahlung die eigentlich verursachenden Störungen und prioritätenbasierten Dispositionsmaßnahmen anteilsmäßig ableiteten. Diese Information kann auch zur anteilsmäßigen Verbuchung von Bonus-Malus-Zahlung zwischen „Schuldigem" und „Betroffenem" der Planabweichung verwendet werden. Für durch Zugstörungen verursachte Anteile ist das jeweilige Verkehrsunternehmen und für Infrastrukturstörungen bzw. Dispositionsmaßnahmen der Infrastrukturbetreiber verantwortlich. Externe Stö-

rungen müssen gesondert berücksichtigt werden. Die konzeptionelle Modellierung kann dem ER-Schema in Abb. 5-5 entnommen werden.

Wie bereits oben genannt, sind für die weitere Betrachtung nur die durch prioritätenbasierte Dispositionsmaßnahmen (D) verursachten Bonus-Malus-Zahlungsanteile $a_vs_disp(BM_x)$ für Verspätungen bzw. $a_avl_disp(BM_x)$ für Anschlußverluste von Bedeutung, da nur diese mittels Prioritätenvergabe beeinflußbar sind. Für diese gilt:

$$a_vs_disp(BM_x) = \sum_{VU_y \in VU(BM_x) \cap D} a_vs(BM_x, VU_y)$$

bzw.

$$a_avl_disp(BM_x) = a_avl(BM_x, ANV) + \sum_{VU_y \in (VU(BM_x) \setminus ANV) \cap D} a_vs(BM_x, VU_y)$$

Wenn man die Gesamtverspätung am Ende des induktiven Verfahrens um $\Sigma\, a_vs(VU_x)$ kürzt, ergibt sich eine theoretische Verspätung für den Fall, daß der Zug mit den gegebenen externen Störfällen mit höchster Priorität durchs Netz gefahren ist und auf keinen Anschluß von anderen Zügen warten mußte.

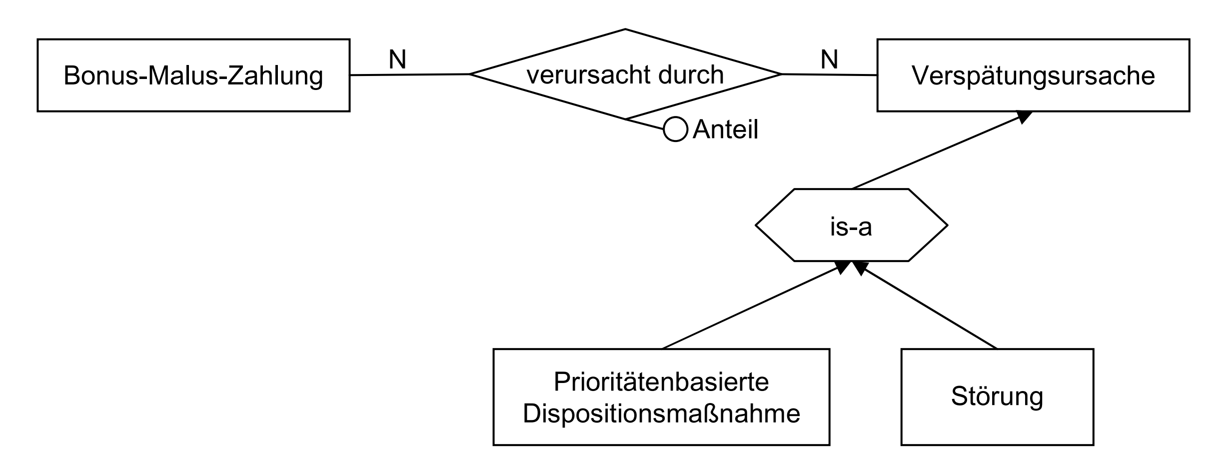


Abb. 5-5: ER-Diagramm Planabweichungsverursachung

5.2.2 Örtliche Zuordnung der Bonus-Malus-Zahlungsverursachungen

Für jede ausgelöste Bonus-Malus-Zahlung eines Zuges aufgrund von Verspätung oder Anschlußverlust kann einerseits der Ort des Auftretens der Abweichung (entsprechende im Bonus-Malus-System festgelegte Verkehrsstation mit Verspätung bzw. Anschlußverlust), andererseits können aber auch relativ leicht aus der im vorigen Abschnitt beschriebenen anteilmäßigen Ursachen der Bonus-Malus-Zahlung die eigentlichen Entstehungsorte der Planabweichungen bestimmt werden, die unmittelbar ursächlich für das Entstehen der Bonus-Malus-Zahlung sind. Dabei werden im folgenden nur die Bonus-Malus-Anteile die durch prioritätenbasierten Dispositionsmaßnahmen verursacht werden (also Anteile in $a_vs_disp(BM_x)$ bzw. $an_avl_disp(BM_x)$, lokal aufgeschlüsselt. Anteile aufgrund von Störungen werden – wegen der Nichtbeeinflußbarkeit durch Prioritäten – hier nicht weiter berücksichtigt.

Als Entstehungsort für Bonus-Malus-Zahlungen sind in diesem Falle Prioritätenabschnitte (für eine spätere Evaluierung von Abschnittsweisen Zugprioritäten) sowie die individuellen Anschlüsse (für eine Evaluierung der Anschlußprioritäten) separat zu betrachten.

Entstehungsorte für Verspätungen sind die vom betroffenen Zug vorher befahrenen Prioritätenabschnitte, in denen prioritätenbasierte Dispositionsmaßnahmen zum Nachteil des betroffenen Zuge aufgrund Abschnittsweiser Zugprioritäten vorgenommen werden, bzw. vorher eingehaltene Anschlüsse, bei denen der Zug als anschlußaufnehmender Zug fungiert. Entstehungsort für Anschlußverluste sind einerseits natürlich der gebrochene Anschluß selbst in Form einer zu geringen Anschlußpriorität, andererseits aber auch die vom Anschlußzubringer vorher befahrenen Prioritätenabschnitte, in denen prioritätenbasierte Dispositionsmaßnahmen aufgrund Abschnittsweiser Zugprioritäten durchgeführt werden, und vorher eingehaltene Anschlüsse, bei denen der Zubringer des gebrochenen Anschlusses als Aufnehmer fungierte.

Die ermittelten Beträge der Bonus-Malus-Zahlungsanteile der entsprechenden Verursacher der Zahlung können einfach auf die zugehörigen Entstehungsorte zugewiesen werden. Bei Anschlußaufnahmen und -brüchen handelt es sich um eine direkte Zuordnung zwischen Ereignis und Ort.

Für die durch einen Anschlußverlust ANV verursachten Bonus-Malus-Kostenanteile $bm_avl_{verurs}(ANV)$ ist nur die Bonus-Malus-Zahlung für den Anschlußbruch BM_x mit Betrag $b(BM_x)$ selbst relevant und es gilt

$$bm_avl_{verurs}(ANV) = a_avl(BM_x, ANV) \cdot b(BM_x)$$

Für durch eine Anschlußaufnahme ANA in Gesamtheit verursachte Bonus-Malus-Kostenanteile $bm_aauf_{verurs}(ANA)$ für Bonus-Malus-Zahlungen BM_1, …, BM_n mit der damit die Anschlußaufnahme repräsentierenden Verspätungsursache $ANA \in VU(BM_i)$ mit i = 1..n gilt

$$bm_aauf_{verurs}(ANA) = \sum_{BM_i, ANA \in VU(BM_i)} a_vs(BM_i, ANA) \cdot b(BM_i)$$

Für Verspätungen des Zuges Z durch prioritätenbasierte Dispositionsmaßnahmen aufgrund Abschnittsweiser Zugprioritäten auf einem Prioritätenabschnitt PA erfolgt zunächst eine Bildung der Menge aller innerhalb des Prioritätenabschnitts entstandenen derartigen Verspätungsursachen mit Z als niederrangigem Zug $VU_pa_nr(Z, PA) = \{VU_1, VU_2, \ldots, VU_n\}$. Darauf basierend errechnet sich die auf PA für Z als niederrangigen Zug verursachten Bonus-Malus-Anteile $bm_pa_nr_{verurs}(Z,PA)$ durch eine Summierung aller verursachten Bonus-Malus-Kostenanteile durch Bonus-Malus-Zahlungen BM_1, BM_2, …, BM_n.

$$bm_pa_nr_{verurs}(Z,PA) = \sum_{BM_i, VU_j \in VU(BM_i) \cap VU_pa_nr(Z,PA)} a_vs(BM_i, VU_j) \cdot b(BM_i)$$

Analog können die auf PA für Z als hochrangigem Zug verursachten Bonus-Malus-Zahlungen $bm_pa_hr_{verurs}(Z, PA)$ bestimmt werden.

Im folgenden Beispiel verkehrt ein relativ niederrangiger Güterzug durch die Prioritätenabschnitte A, B und C. Er erhält in Abschnitt A 45 Minuten Verspätung durch eine Signalstörung und weitere 30 Minuten durch Überholvorgänge, in der zwischen A und B gelegenen Verkehrsstation weitere 15 Minuten durch Warten auf einen Anschluß (Aufnahme eines Güterwagens), in Abschnitt B 45 Minuten durch diverse Überholvorgänge und in Abschnitt C weitere 45 Minuten durch Überholvorgänge. Im Bahnhof C wird nun ein Anschluß auf einen weiteren Güterzug gebrochen. Die maximale Wartezeit ist auf 30 Minuten festgelegt, der Anschlußbruch erfolgt ab einer Verspätung von zwei Stunden. Für den Anschlußbruch wurde eine Bonus-Malus-Zahlung von 2.000 EUR festgelegt.

Die Verspätung des Zubringerzuges beträgt am Endbahnhof von C 180 Minuten. Somit werden (30 : 180) · 2.000 EUR = 333,33 EUR Bonus-Malus-Anteil auf die eigentliche (zu geringe) Anschlußpriorität angerechnet. Die restlichen 1.666,67 EUR werden mit dem induktiven Verfahren Vorfällen auf dem Zuglaufweg des Zubringerzuges zugeordnet. Da nach dem Passieren von Abschnitt B die Verspätung durchgängig über zwei Stunden beträgt, wird das Verfahren nach dem Passieren von B abgebrochen und die Zahlung nur auf das (nicht weitere betrachtete) externe Ereignis, die Abschnitte A und B sowie den eingehaltenen Anschluß umgelegt. Da zwischenzeitlich keine Verspätungen aufgeholt werden, kann eine einfache relative Aufteilung der Zahlung erfolgen. Die relevante Gesamtverspätung bis zum Passieren von B beträgt 45 + 30 + 15 + 45 Minuten = 135 Minuten. Hierbei erhält A einen Anteil von (30 : 135) · 1.666,67 EUR = 370,37 EUR der Anschluß einen Anteil von (15 : 135) · 1.666,67 EUR = 185,19 EUR und B einen Anteil von (45 : 135) * 1.666,67 EUR = 555,56 EUR.

An dieser Stelle soll noch einmal wiederholt werden, daß es hier lediglich um eine „neutrale“ (unabhängig vom Verursacher der eigentlichen Störungen) anteilmäßige Umlegung von Bonus-Malus-Zahlungen auf teilweise verursachende prioritätenbasierte Dispositionsmaßnahmen handelt und diese Beträge nicht unbedingt reale Zahlungen zwischen den beteiligten Parteien darstellen müssen. Hierfür ist eine detaillierte Analyse der primären Störungen erforderlich, die für das hier vorgeschlagene Verfahren wegen der Neutralitätsanforderung explizit unterlassen werden soll.

5.2.3 Bonus-Malus-Verursachungen als Grundlage für die Bewertung und Anpassung von Prioritäten

Aufgrund der lokalen Zuweisung verursachter Bonus-Malus-Zahlungsanteile kann eine nähere Bewertung der Abschnittsweisen Zugprioritäten auf betroffenen Prioritätenabschnitten sowie der Anschlußprioritäten betroffener Anschlüsse erfolgen. Hierzu wird eine Zuordnung in zweierlei Weise vorgenommen. Einerseits soll zu jedem relativ hochrangigen Zug auf einem Prioritätenabschnitt bzw. eingehaltenen Anschluß erfaßt werden, inwieweit dieser durch Behinderung niederrangiger bzw. anschlußaufnehmender Züge anteilsmäßig prioritätenbedingte Bonus-Malus-Zahlungsanteile desselben verursacht hat (hierbei handelt es sich um $bm_pa_hr_{verurs}(Z, PA)$ bzw. $bm_aauf_{verurs}(ANA)$). Eine Häufung dieser Vorfälle für einen Modellzug bzw. Anschluß ist ein Indiz für eine möglicherweise zu hohe Abschnittsweise Zugpriorität auf diesem Prioritätenabschnitt bzw. eine zu hohe Anschlußpriorität für den einge-

haltenen Anschluß. Andererseits soll auch für den relativ niederrangigen Zug bzw. nicht eingehaltenen Anschluß protokolliert werden, inwieweit dieser auf Prioritätenabschnitten seitens Behinderungen höherrangiger Züge bzw. durch nicht eingehaltene Anschlußaufnahmen anteilsmäßig prioritätenbedingte Bonus-Malus-Zahlungsanteile verursacht hat (hierbei handelt es sich um $bm_pa_nr_{verurs}(Z, PA)$ bzw. $bm_avl_{verurs}(ANV)$). Eine Häufung dieser Vorfälle für einen Modellzug bzw. Anschluß ist ein Indiz für eine möglicherweise zu niedrige Abschnittsweise Zugpriorität auf diesem Prioritätenabschnitt bzw. zu niedrige Anschlußpriorität für den gebrochenen Anschluß. Schließlich soll für jeden Zug zusätzlich zu anteilsmäßig verursachten prioritätenbedingten Bonus-Malus-Zahlungsanteilen der jeweils aktuell maßgebliche Spezielle Betriebszustand in diesem Prioritätenabschnitt bzw. bei dem entsprechenden Anschluß gespeichert werden, um eine Differenzierung zwischen Standard- und Störfallprioritäten vornehmen zu können.

Sowohl für die relativ hochrangigen behindernden Modellzüge als auch für die relativ niederrangigen, für die eingehaltenen wie auch für die gebrochenen Anschlüsse, können Toleranzgrenzen festgelegt werden, wie hoch die Summen der jeweils entstehenden Bonus-Malus-Zahlungen für einen bestimmten Modellzug in einem Prioritätenabschnitt oder Anschluß unter Standardbedingungen bzw. Speziellen Betriebszuständen in einer vorgegebenen Anzahl von Betriebstagen sein dürfen, ohne daß eine sofortige Prioritätenoptimierung ausgelöst werden soll. Damit sollen kurzfristige und häufige Feinjustierungen der Prioritäten vermieden werden. Der Fokus soll auf der Lösung längerfristiger Probleme liegen.

Diese Erweiterung ist in Abb. 5-6 bzw. Abb. 5-7 als ER-Diagramm dargestellt. Für in Prioritätenabschnitten verursachte Bonus-Malus-Zahlungen werden die Toleranzgrenzen individuell für jeden Modellzug (unter Geltung von Standard- und Störfallprioritäten) als relativ niederrangiger bzw. relativ hochrangiger Zug definiert. Als Betriebstage werden dann aber nur die Betriebstage berücksichtigt, in denen der Modellzug im Prioritätenabschnitt unter dem ggf. angegebenen Speziellen Betriebszustandstyp gefahren ist. Alternativ können auch mehrere, derselben Störfallpriorität assoziierte Spezielle Betriebszustandstypen in einer Toleranzgrenze zusammengefaßt werden, wobei hier aus Gründen vereinfachender Darstellung darauf verzichtet werden soll. Es besteht die Möglichkeit, im Falle von Aktualisierungen ein Reset-Datum festzulegen, ab welchem Betriebstag die Toleranzgrenzen überwacht werden. Entsprechend wird bei Anschlüssen für eingehaltene und gebrochene Anschlüsse verfahren.

Wenn eine derartige Toleranzgrenze überschritten wird, erfolgt eine Anpassung der Dispositionsprioritäten. Es gibt es folgende Möglichkeiten:

- Bei der Toleranzgrenzenüberschreitung für durch einen relativ hochrangigen (niederrangigen) Modellzug auf einem Prioritätenabschnitt verursachte Bonus-Malus-Zahlungen wird die entsprechende Abschnittsweise Zugpriorität reduziert (erhöht). Um eine gewisse längerfristige Monotonie zu erhalten, sollen die Prioritäten nur leicht verändert werden. Bei einer Reduzierung soll die neue Priorität nur eine Rangstufe unterhalb des höchstrangigen Modellzuges, welcher durch den betroffenen Modellzug behindert wurde, gesetzt werden. Bei einer Erhöhung der Priorität gilt selbiges entsprechend.

- Bei der Toleranzgrenzenüberschreitung für durch einen eingehaltenen (gebrochenen) Anschluß verursachte Bonus-Malus-Zahlungen wird die entsprechende Anschlußpriorität reduziert (erhöht). Hier soll wiederum nur eine leichte Veränderung erfolgen, z. B. kann eine Veränderung um 10% der aktuell vorgegebenen maximalen Wartezeit erfolgen.

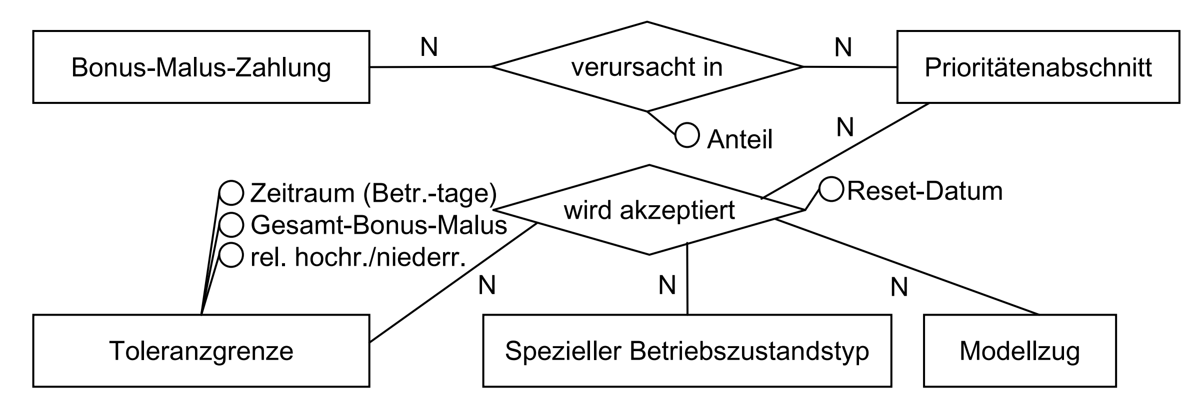


Abb. 5-6: ER-Diagramm Verursacher von in Prioritätenabschnitten verursachten Bonus-Malus-Zahlungen

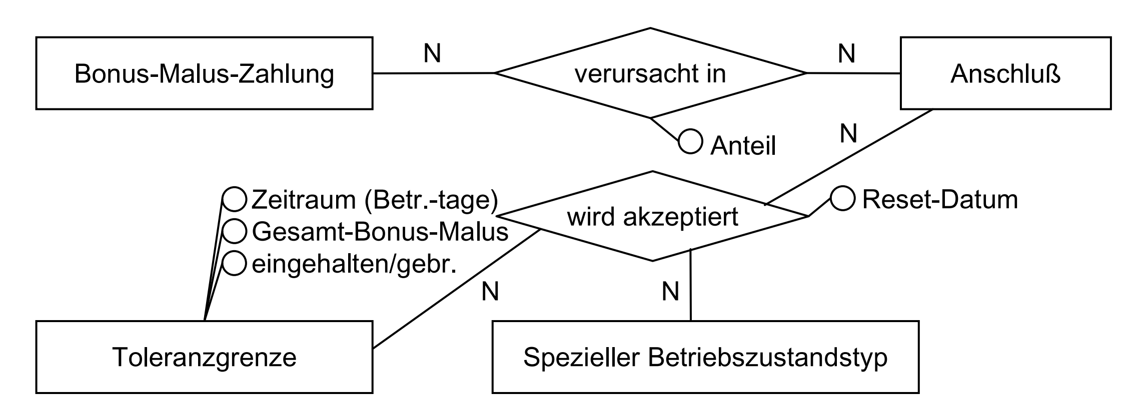


Abb. 5-7: ER-Diagramm Verursacher von durch abgewartete bzw. gebrochene Anschlüsse verursachte Bonus-Malus-Zahlungen

5.3 Gewährleistung einer langfristigen Optimierung

Im folgenden soll die Frage diskutiert werden, wie sichergestellt werden kann, daß im Rahmen einer Prioritätenoptimierung ein wirklicher Fortschritt im Sinne einer Verbesserung erreicht wird. Wenn dies nicht sichergestellt werden kann, besteht die Gefahr eines gegenseitigen „Hochschaukelns“ von Prioritäten (Prioritätenzyklen), ohne dabei dem Optimierungsziel wirklich näherzukommen.

Hierzu sollten die Auswirkungen von Dispositionsprioritäten in Form von durch konkrete Dispositionsmaßnahmen verursachten Bonus-Malus-Zahlungsanteilen längerfristig für spätere Vergleiche zwischen verschiedenen bereits „getesteten“ Konstellationen verwendet werden. Sollte das in 5.2 beschriebene Bewertungsverfahren eine Anpassung von Standard- oder Störfallprioritäten vorschlagen, aus der eine Konstellation resultiert, die bereits in der Vergangenheit getestet wurde und bereits als ungünstig bewertet wurde, soll statt dessen – falls

möglich – durch Wahl einer alternativen Priorität eine bereits als besser getestete Konstellation erreicht werden. Dadurch wird eine Wiederholung von bereits überprüften und als unzureichend qualifizierten Prioritätenkonstellationen vermieden und folglich mittelfristig ein wirklicher Fortschritt im Optimierungsprozeß sichergestellt.

Konkret wird dabei wieder zwischen Anschlußprioritäten und Abschnittsweisen Zugprioritäten unterschieden. Für erstere ist das Verfahren relativ einfach durchführbar. An jedem Betriebstag, an dem der Anschluß vorgesehen ist, kann dieser entweder eingehalten oder gebrochen werden. In beiden Fällen kann eine anteilsmäßige Verursachung einer Bonus-Malus-Zahlung erfolgen, die entsprechend protokolliert werden muß. Man kann den durchschnittlich durch einen Anschluß verursachten Bonus-Malus-Zahlungsanteil abhängig von der konkret verwendeten Anschlußpriorität bestimmen. Hierbei müssen getrennte Statistiken für Standard- und Störfallprioritäten geführt werden. Entsprechend läßt sich für die Standard- und jede Störfallpriorität eine Ordnung der möglichen Anschlußprioritätsbeträge bzgl. der verursachten Bonus-Malus-Zahlungsanteile erstellen. Falls nun im Optimierungsverfahren eine entsprechend ungünstige Anschlußpriorität als Standard- oder Störfallpriorität vorgeschlagen wird, kann statt dessen der entsprechend dieser Ordnung optimale Wert vergeben werden. Dabei ist zu beachten, daß aufgrund der statistischen Sicherheit eine bestimmte Mindestanzahl von Betriebstagen überprüft sein sollte, bevor Aussagen bzgl. der durchschnittlich verursachten Bonus-Malus-Zahlungsanteile für eine konkrete Priorität erfolgen können.

Dementsprechend werden für den Anschluß AN, den maßgeblichen Speziellen Betriebszustandstyp SBZ und die Anschlußpriorität PR in zwei Funktionen

$$\text{bt_anschl}(AN, SBZ, PR) \text{ und}$$
$$\text{avg_bm_anschl}_{\text{verurs}}(AN, SBZ, PR)$$

die Anzahl der Betriebstage unter der die Konstellation gültig war und der durchschnittlich durch diese Anschlußpriorität wegen Einhalten und Bruchs verursachte Bonus-Malus-Anteil zugewiesen. Sobald bt_anschl(AN, SBZ, PR) größer als eine vorgegebene untere Schranke ist, kann die Priorität PR entsprechend in eine Ordnungsrelation

$$\text{order_anschl_pr}(AN, SBZ) \rightarrow [PR_1, PR_2, \ldots, PR_n]$$

eingefügt werden, auf die bei künftigen Optimierungen zurückgegriffen werden kann.

Für Abschnittsweise Zugprioritäten ist das Verfahren wegen der Wechselwirkungen mit anderen Zügen etwas komplizierter. Hier müssen entsprechende Statistiken für alle Modellzugpaare aufgestellt werden, zwischen denen ein mittels einer prioritätenbasierten Dispositionsmaßnahme aufgelöster Belegungskonflikt auftritt. Einem Modellzug MZ auf einem Prioritätenabschnitt PA sind so eine Menge $ZP(MZ,PA) = \{MZ_1, MZ_2, \ldots, MZ_n\}$ möglicher Konfliktzüge zugeordnet. Hierbei soll ein Mindestwert für tatsächliche aufgetretene Belegungskonflikte festgelegt werden, so daß eine nennenswerte Wahrscheinlichkeit für zukünftige Belegungskonflikte für die Modellzüge in ZP(MZ,PA) zu erwarten ist. Ferner muß weiterhin jede zum Zeitpunkt einer Dispositionsmaßnahme gültige Kombination aus Standard-

und Störfallpriorität beider Züge gesondert untersucht werden. Die Absolutbeträge der beiden Abschnittsweisen Zugprioritäten sind dabei nicht von entscheidender Bedeutung, vielmehr ist deren relative Ordnung (größer, kleiner oder gleich) entscheidend. Für jede dieser relativen Prioritätenordnungen einer Kombination von Standard- bzw. Störfallprioritäten soll nun der mittlere verursachte Bonus-Malus-Zahlungsanteil pro Betriebstag durch Konflikte dieser beiden Modellzüge errechnet werden. Man beachte, daß ein Bonus-Malus-Anteil nur im Falle eines tatsächlichen Belegungskonfliktes zweier Züge eintritt.

Entsprechend wird für den Prioritätenabschnitt PA und die Modellzüge MZ_1 und MZ_2 mit den maßgeblichen Speziellen Betriebszustandstypen $SBZ_1 \in SBZ(MZ_1,PA)$ und $SBZ_2 \in SBZ(MZ_2,PA)$ und der relativen Ordnung der Abschnittsweisen Zugprioritäten RO in Funktionen

$$bt_pa(PA, MZ_1, SBZ_1, MZ_2, SBZ_2, RO) \text{ und}$$
$$avg_bm_pa_{verurs}(PA, MZ_1, SBZ_1, MZ_2, SBZ_2, RO)$$

wiederum die Anzahl der Betriebstage der Konstellation und der mittlere pro Betriebstag als niederrangiger oder hochrangiger Zug verursachte Bonus-Malus-Anteil zugewiesen. Ein nicht vorhandener Belegungskonflikt unter der vorgegebenen Konstellation geht in die Statistik gewichtend mit einem Bonus-Malus-Anteil von 0 für diesen Betriebstag ein.

Damit sind nun Vergleiche möglich, die sich auf tatsächlich zu erwartende Betriebskonstellationen fokussieren. Sobald $bt_pa(PA, MZ_1, SBZ_1, MZ_2, SBZ_2, RO)$ eine bestimmte Mindestanzahl (statistische Sicherheit!) überschreitet, kann entsprechend der berechneten Erwartungswerte bestimmt werden, welche relative Ordnung (größer, kleiner, gleich) „besser“ als die andere ist. Es ergibt sich die Ordnungsrelation

$$order_pa_pr(PA, MZ_1, SBZ_1, MZ_2, SBZ_2) \rightarrow [RO_1, RO_2, RO_3]$$

Diese Ergebnisse können nun bei der Anpassung von Abschnittsweisen Zugprioritäten durch das Optimierungsverfahren verwendet werden, um eine Anpassung auf eine bereits als ungünstig getestete relative Konstellation zu vermeiden. Dies soll am Beispiel eines Modellzuges A verdeutlicht werden, dessen Standardpriorität infolge einer Toleranzgrenzenüberschreitung durch das Optimierungsverfahren verändert werden soll. Hierbei müssen alle Betriebstage berücksichtigt werden, bei denen A mit Standardpriorität verkehrte. Die Anzahl dieser Betriebstage wird mit BT bezeichnet. Für mögliche Konfliktzüge aus ZP(A,PA) spielt es keine Rolle, ob diese mit Standard- oder einer Störfallpriorität verkehrten. Sämtliche Konstellationen mit einer entsprechenden Mindestanzahl an Betriebstagen sollen entsprechend ihrer relativen Häufigkeit berücksichtigt werden, die sich folgendermaßen bestimmen lassen:

$$rel_h(PA, A, -, MZ_2, SBZ_2) = \frac{\sum_{RO} bt_pa(PA, A, -, MZ_2, SBZ_2, RO)}{BT}$$

Für die zu testende neue Standardpriorität können einfach die daraus resultierenden relativen Ordnungen zu den derzeit vorgesehenen Standard- und Störfallprioritäten der möglichen Konfliktzüge bestimmt werden, so daß die entsprechenden Erwartungswerte der verursachten Bonus-Malus-Zahlungsanteile summiert werden können. Daraus ergibt sich der

täglich von Modellzug A mit der neuen Standardpriorität PR zu erwartende Bonus-Malus-Zahlungsanteil pro Betriebstag wobei die gemäß den derzeit vorgesehenen Prioritäten relativen Ordnungen RO entsprechend eingesetzt werden:

$$avg_total_bm_pa_{verurs}(PA,A,-,PR) = \sum_{MZ_2 \in ZP(A), SBZ_2 \in SBZ(MZ_2,PA)} avg_bm_pa_{verurs}(PA,A,-,MZ_2,SBZ_2,RO) \cdot rel_h(PA,A,-,MZ_2,SBZ_2)$$

Es wird überprüft, ob eine andere in der Vergangenheit getestete Standardpriorität PR_x für Modellzug A mit dieser Methodik einen niedrigeren zu erwartenden Bonus-Malus-Zahlungsanteil $avg_total_bm_pa_{verurs}(PA,A,-,PR_x)$ erreicht. In diesem Fall soll die Standardpriorität auf diesen sich bereits als günstiger erwiesenen Wert PR_x gesetzt werden. Damit wird ein wiederholtes Austesten von sich in der Vergangenheit als ungünstig erwiesenen Werten vermieden.

Da die Anpassung der Abschnittsweisen Zugpriorität auch Auswirkungen auf die relative Priorität der an prioritätenbasierten Dispositionsmaßnahmen mit A beteiligten Modellzüge hat, sollen nach einer derartigen Optimierung neben der Priorität von A sämtliche Prioritäten der möglichen Konfliktzüge eine gewisse Zeit festgeschrieben (geschützt) bleiben, um die Auswirkungen dieser Änderung geeignet bewerten zu können und damit sich nicht verschiedene kurz hintereinander durchgeführte Optimierungsschritte gegenseitig beeinflussen bzw. aufheben. Die Messung von Toleranzgrenzenüberschreitungen beginnt zu diesem Zeitpunkt von vorn (Verwendung des in Abschnitt 5.2.3 eingeführten Reset-Datums). Für nicht möglicherweise in Konflikt mit A stehende Modellzüge bleiben hingegen die Bonus-Malus-Ereignisse erhalten und es kann jederzeit bei einer eintretenden Toleranzgrenzenüberschreitung eine erneute Optimierung erfolgen.

Entsprechend muß das Verfahren zur Reaktion auf Toleranzgrenzenüberschreitungen noch einmal genauer spezifiziert werden. Mehrere gleichzeitig eintreffende Grenzüberschreitungen werden entsprechend in der absteigenden Ordnung der Gesamthöhe der absolut verursachten Bonus-Malus-Zahlungsanteile abgehandelt. Dadurch wird sichergestellt, daß bei zwei gleichzeitigen Toleranzgrenzenüberschreitungen von miteinander in Wechselwirkung stehenden Prioritäten nur eine Reaktion durch das Verfahren erfolgt, nämlich auf das „gravierendere“ Problem. Durch den daraus folgenden automatischen Neubeginn der Beobachtung von Toleranzgrenzenüberschreitungen (Reset) für das zweite Problem, „verschwindet“ diese Überschreitung und es erfolgt keine weitere – möglicherweise die erste Optimierung koterkarierende – Feinjustierung. Bei größeren Änderungen des Betriebsprogramms ist eine Löschung der alten Vergleichswerte erforderlich, da dann die Vergleichbarkeit nicht mehr gegeben ist.

Da die Optimierung der Prioritäten grundsätzlich jeweils am eigentlichen Entstehungsort der ungünstigen Bonus-Malus-Zahlungen erfolgten, und sich zwischen den einzelnen Prioritätenabschnitten per Definition ausreichende Puffermöglichkeiten für Züge befinden, kann man (bei Betrachtung eines längeren Zeitraumes) weitgehende gegenseitige Unabhängigkeit und Nichtbeeinflussung von Prioritäten auf verschiedenen Prioritätenabschnitten an-

nehmen. Durch die Prioritätenvergabe in einem Prioritätenabschnitt werden keine Bonus-Malus-Zahlungen in anderen Abschnitten verursacht, sondern nur im eigenen Abschnitt. Daher kann man das globale Optimum der verursachten Bonus-Malus-Zahlungsanteile aus der „Summe“ aller lokalen Optima in den einzelnen Prioritätenabschnitten entnehmen.

Die berechneten, auf Prioritätenabschnitte umgelegten lokalen Bonus-Malus-Anteile können ggf. auch unterstützend zur Infrastrukturoptimierung verwendet werden. Wenn auf einem entsprechenden Abschnitt durch Optimierung kein nachhaltiger Fortschritt bzgl. ungünstiger Bonus-Malus-Zahlungsanteile erreicht werden kann, dann deutet das auf einen infrastrukturellen Engpaß hin, der ggf. nur durch weitere oder verbesserte Überholungsmöglichkeiten in diesem Prioritätenabschnitt entlastet werden kann. Gleiches gilt analog für minimale Bonus-Malus-Zahlungsanteile auf einem Prioritätenabschnitt. Hier stellt sich ggf. die Frage nach einer überdimensionierten Infrastruktur. Eine genaue Analyse darf sich aber wegen der starken Fahrplanabhängigkeit nicht nur auf diese Aussagen verlassen, vielmehr müssen ggf. auch fahrplanunabhängige Untersuchungen ergänzend durchgeführt werden, insbesondere dann, wenn mit größeren Änderungen des Betriebsprogramms zu rechnen ist. Details zu dieser Thematik können im Rahmen dieser Arbeit leider nicht behandelt werden.

6 Entwurf eines Prioritätenvergabe- und -optimierungssystems

Im folgenden Kapitel soll nun der Entwurf des zur Umsetzung dieses im vorhergehenden Kapitel dargestellten Verfahrens angedachte Systems dokumentiert werden. Dazu soll zunächst die Architektur des Systems und die Einbettung in eine Anwendungsumgebung dargestellt werden. Anschließend erfolgt ein Abriß der Nutzung deduktiver und aktiver Komponenten innerhalb dieser Architektur. Schließlich sollen die Funktionalität des Systems detailliert beschrieben, sowie mögliche Regelsätze des Systems diskutiert werden.

6.1 Architektur und Eingliederung in eine Systemumgebung

Um effektiv eingesetzt werden zu können, muß das zu entwerfende Prioritätenvergabe- und -optimierungssystem in eine Systemumgebung für die prioritätenbasierte Eisenbahnbetriebsdisposition integriert werden. Für diese Gesamtarchitektur werden hier drei voneinander getrennte Teilsysteme vorgeschlagen, die den Eisenbahnbetrieb in der „Außenwelt" überwachen und koordinieren. Dabei handelt es sich um ein prioritätenbasiertes Dispositionssystem, ein Qualitätsüberwachungssystem und das eigentliche Prioritätenvergabe- und -optimierungssystem (vgl. Abb. 6-1). Sie sind über Schnittstellen miteinander verbunden.

Die **„Außenwelt"** repräsentiert den tatsächlich stattfindenden Eisenbahnbetrieb. Auftretende elementare Störungen an Fahrzeugen oder Fahrweg bzw. organisatorische Mißstände können in Belegungs-, Anschluß- oder Umlaufkonflikten resultieren, die im Rahmen der Disposition aufgelöst werden müssen. Vorab definierte Typen von Verkehrsmeldungen (insbesondere automatische oder manuelle Zuglaufmeldungen) werden in der „Außenwelt" generiert und signalisiert.

Das **prioritätenbasierte Dispositionssystem** dient der eigentlichen Konflikterkennung und -lösung auf Basis der vom Prioritätenvergabe- und -optimierungssystem vorgegebenen Prioritäten. Die Resultate der Disposition werden der „Außenwelt" zur Umsetzung mitgeteilt. Außerdem leitet das Dispositionssystem aufgetretene, für die Prioritätenvergabe unter Umständen relevante Störungsmeldungen an das Prioritätenvergabe- und -optimierungssystem weiter.

Zusätzlich existiert im Gesamtkontext ein **Qualitätsüberwachungssystem** zur Bewertung des Betriebs anhand des vorgegebenen Bonus-Malus-Systems. Diesbezügliche standardisierte Beobachtungen werden wiederum dem Prioritätenvergabe- und -optimierungssystem zur evtl. erforderlichen Prioritätenanpassung mitgeteilt. Die Prioritäten werden durch das Prioritätenvergabe- und -optimierungssystem aufgrund der übermittelten Eingabeereignisse generiert und dem Dispositionssystem rechtzeitig zur Verfügung gestellt.

Ein Vorteil einer Realisierung mit mehreren separaten Teilsystemen anstelle eines integrierten Gesamtsystems besteht unter anderem darin, daß die Systeme leichter wartbar sind. Insbesondere Dispositionssysteme sind in der Regel sehr voluminöse Systeme, die einen großen Aufwand für Anpassungen benötigen, so daß sich die Auslagerung von nicht zur Kernapplikation gehörigen Prozessen in separate Systeme empfiehlt. Ferner besteht zwi-

schen allen drei Teilsystemen nur eine geringe zeitliche und organisatorische Synchronisation der Prozesse, so daß eine getrennte Architektur mit Schnittstellen gegenüber einem integrierten Ansatz zu bevorzugen ist. Außerdem ist das Dispositionssystem nicht zwangsläufig an die exakte Umsetzung der vorgeschlagenen Prioritäten gebunden, es hat die alleinige Entscheidungsgewalt. Das Prioritätenvergabe- und -optimierungssystem kann keinen direkten Einfluß auf spezifische Dispositionsmaßnahmen nehmen. Durch eine klare Definition von Schnittstellen ist darüber hinaus ein Austausch von einzelnen Teilsystemen unter Beibehaltung der übrigen Systeme relativ einfach zu bewerkstelligen, wohingegen dies bei einem integrierten Gesamtsystem praktisch unmöglich ist.

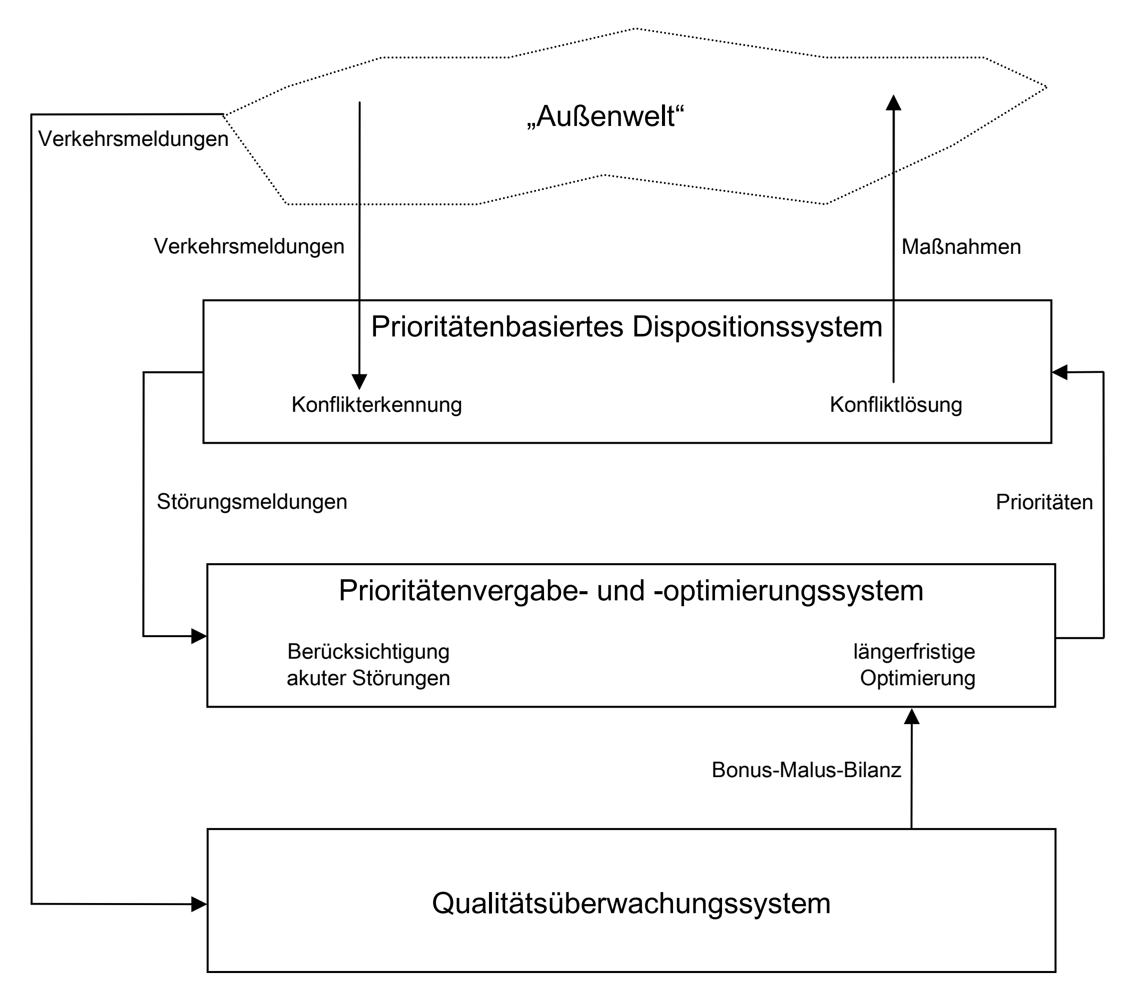


Abb. 6-1: Einordnung in Gesamtarchitektur

Im folgenden soll eine Architektur des eigentlichen Prioritätenvergabe- und -optimierungssystems aufgezeigt werden. Anschließend erfolgt eine Darstellung der wesentlichen Anforderungen an das in die Gesamtumgebung zu integrierende prioritätenbasierte Dispositionssystem sowie des Qualitätsüberwachungssystems.

6.1.1 Architektur des Prioritätenvergabe- und -optimierungssystems

Die bereits ausführlich beschriebene Kernaufgabe des eigentlichen Prioritätenvergabe- und -optimierungssystems besteht darin, dem Dispositionssystem zur Konfliktlösung eine Entscheidungsunterstützung in Form von möglichst im Sinne eines vorgegebenen Bonus-Malus-Systems verkehrswirtschaftlich günstigen, zulässigen Abschnittsweisen Zug- und Anschlußprioritäten für die Disposition zu liefern. Sie haben für das Dispositionssystem empfehlenden Charakter. Die konkrete Umsetzung der gelieferten Prioritäten in konkrete Dispositionsentscheidungen verbleibt beim Disponenten.

Die Akzeptanz eines Prioritätenvergabe- und -optimierungssystems ist bei den Anwendern nur gegeben, wenn dadurch insbesondere im Zusammenhang mit der Störungsbehandlung keine nennenswerte zeitliche Verzögerung bei der Disposition resultiert. Daher soll die Übermittlung der Abschnittsweisen Zug- und Anschlußprioritäten unabhängig von der eigentlichen Betriebsdisposition und i. d. R. in einem ziemlich kurzen Zeitraum auf Anforderung des Dispositionssystems erfolgen.

Im Prioritätenvergabe- und -optimierungssystem sollen die bei der Disposition zu verwendenden Abschnittsweisen Zug- und Anschlußprioritäten für Standard- und Störfälle jederzeit abrufbar sein und „im Hintergrund" aufgrund der anhand des Maßstabs des Bonus-Malus-Systems erreichten Ergebnisse langfristig weiter optimiert werden. Die Kommunikation zwischen den Komponenten des zu entwickelnden Prioritätenvergabe- und -optimierungssystem ist schematisch in Abb. 6-2 dargestellt:

Das Prioritätenvergabe- und -optimierungssystem besteht somit aus einer Benutzerschnittstelle und den fünf zentralen Komponenten Administrator der Speziellen Betriebszustände (Störfälle), Prioritätenvergabekomponente, Toleranzgrenzenüberprüfer für Bonus-Malus-Zahlungen, Prioritätenoptimierungs- und Prioritäteninitialisierungskomponente. Während die vier erstgenannten Komponenten permanent verwendet werden, wird die (partielle und optionale) Prioritäteninitialisierung ausschließlich bei Änderungen von wesentlichen Eingangsdaten eingesetzt.

Das Prioritätenvergabe- und -optimierungssystem tritt mit dem Dispositionssystem und dem Qualitätsüberwachungssystem über definierte Schnittstellen in Interaktion. Außerdem existiert die bereits genannte Benutzerschnittstelle. Diese Schnittstellen sollen jetzt noch kurz im einzelnen diskutiert werden.

Die Schnittstelle zum Dispositionssystem ist entscheidend für das Gesamtkonzept zur Zugdisposition mit Hilfe von optimierten Prioritäten. Sie wird in beiden Richtungen benötigt, zum einen zur Signalisierung von Betriebsereignissen (Initialisierung/Terminierung/Aktualisierung von Speziellen Betriebszuständen), zum anderen auf Anforderung des Dispositionssystems zur Rückspielung von Abschnittsweisen Zug- und Anschlußprioritäten in das Dispositionssystem als Vorgabe für Konfliktlösungen.

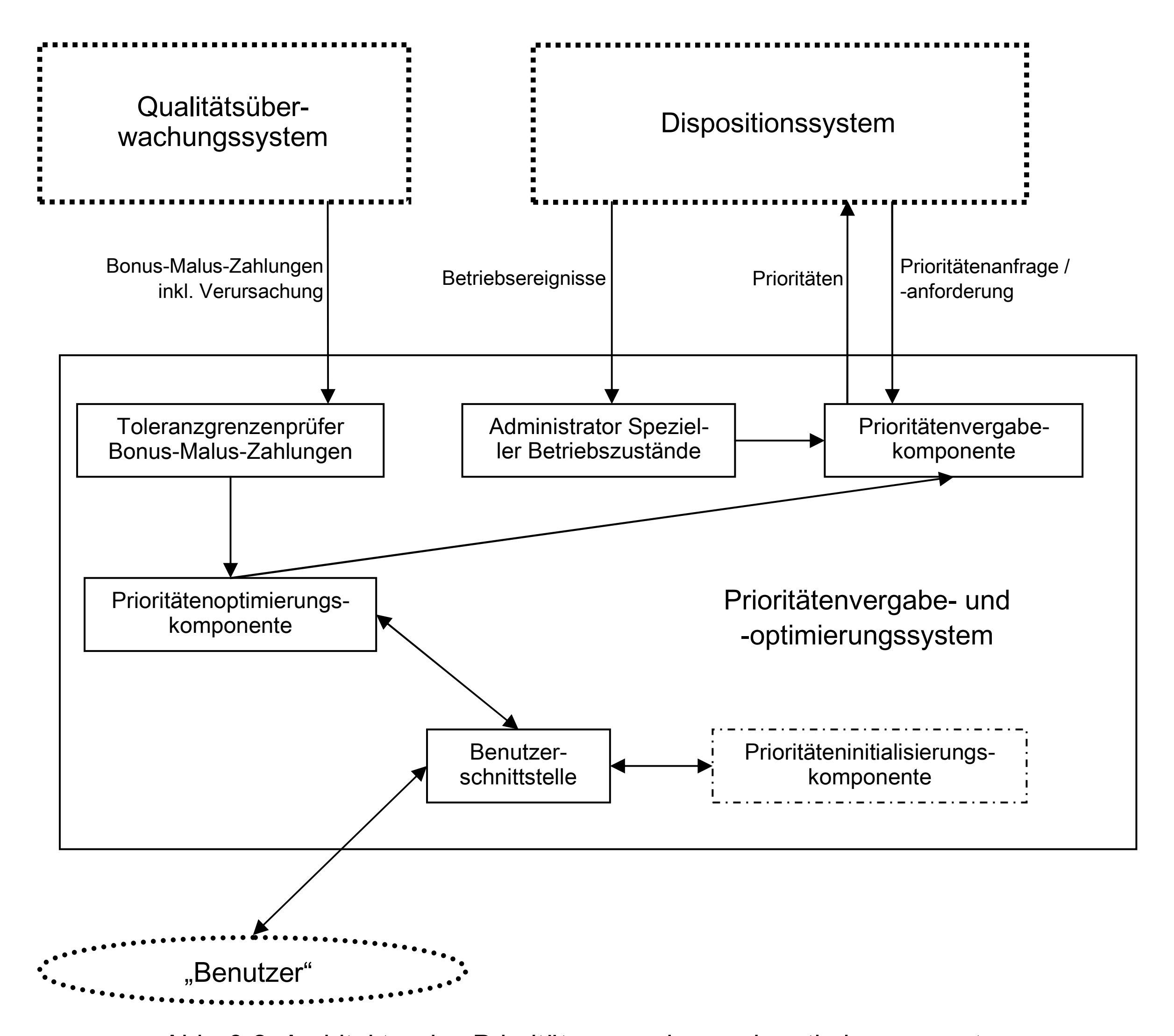


Abb. 6-2: Architektur des Prioritätenvergabe- und -optimierungssystems

Die Ereignisse und Prioritäten werden entsprechend der in Kapitel 4 dargestellten Modelle codiert, so daß eine geeignete Verarbeitung in den jeweiligen Systemen ermöglicht wird. Signalisierte Betriebsereignisse haben dementsprechend eine eindeutige ID und einen vom Dispositionssystem vergebenen Zeitstempel. Es handelt es im Detail um folgende im Zusammenhang mit Störfallprioritäten relevanten Ereignisse (vgl. 4.3):

- Planabweichungsmeldungen,
- Störungsereignisse zu punktuellen Zugstörungen
- Störungsereignisse zu kontinuierlichen Zugstörungen und
- Störungsereignisse zu kontinuierlichen Infrastrukturstörungen.

Die Schnittstelle zwischen Qualitätsüberwachungssystem und Prioritätenvergabe- und -optimierungssystem dient dazu, die vom Qualitätsüberwachungssystem detektierten Bonus-Malus-Zahlungen inkl. der Prioritätenabschnitte und Anschlüsse, die als eigentliche Verursacher der Zahlung gelten, als Ereignisbotschaften an das Prioritätenvergabe- und -optimierungssystem zwecks Prüfung der vorgegebenen Toleranzgrenzen und ggf. anschließender Prioritätenoptimierung zu übergeben. Die Übermittlung kann unmittelbar bei Detektion oder

auch verzögert erfolgen. Datentransfer vom Prioritätenvergabe- und -optimierungssystem in das Qualitätsüberwachungssystem findet nicht statt.

Die Benutzerschnittstelle dient dazu, das System im Betrieb zu nehmen und Aktualisierungen (z. B. an Fahrplänen und normativen Vorgaben) vornehmen zu können, was entsprechende (partielle) Prioritäteninitialisierungen zur Folge haben kann. Ferner kann mit Hilfe der Benutzerschnittstelle durch den Benutzer Einfluß auf die Prioritätenoptimierung genommen werden. So besteht die Möglichkeit, bestimmte „kritische“ Prioritätenoptimierungen unter dem Vorbehalt einer manuellen Freigabe („commit“) durch den Benutzer zu stellen oder bestimmte Prioritäten zu sperren. Letztlich kann die Benutzerschnittstelle auch zur Information des Benutzers z. B. bei nicht möglicher Optimierung verwendet werden.

6.1.2 Anforderungen an das zugehörige prioritätenbasierte Dispositionssystem

Ein in das hier entworfene Gesamtsystem integrierbares Dispositionssystem dient der automatischen bzw. manuellen rechnerunterstützten Feinkonstruktion von konfliktfreien Trassen bei Betriebsdisposition auf Basis von Wunschfahrlagen (ursprünglicher Sollfahrplan), vorgegebenen Prioritäten und Meldungen aus der „Außenwelt“. Hierbei ist es denkbar, bereits auf dem Markt bestehende prioritätenbasierte Systeme mit entsprechend nachzurüstenden Schnittstellen zu verwenden. Welche existierenden Systeme genau für eine Einbindung geeignet sind, soll ebenfalls im weiteren Verlauf dieses Abschnittes erarbeitet werden. Auch die Neuentwicklung eines eigenständigen Dispositionssystems ist alternativ möglich.

Das benötigte Dispositionssystem arbeitet über einen längeren Zeitraum hinweg, wobei die Ergebnisse evtl. nur über einen kurzen, unmittelbar bevorstehenden Zeitraum hinweg gültig sind. Die eigentliche Kernaufgabe des Dispositionssystems besteht in der Konflikterkennung und -lösung. Als mögliche Konflikte sind Belegungs-, Anschluß- und Umlaufkonflikte möglich. Eine temporäre Sperrung eines Gleisabschnitts wird vom System in diesem Zusammenhang wie ein Belegungskonflikt eines strikt nachrangigen Zuges behandelt. Erstere werden mit Hilfe der Abschnittsweiten Zugprioritäten, letztere mit Hilfe von Anschlußprioritäten entschieden.

Neben Sperrungen sollte das Dispositionssystem auch auf weitere aus der „Außenwelt“ gemeldete temporär eingeschränkte Ressourcen auf Fahrweg- oder Fahrzeugseite reagieren (Geschwindigkeitsbeschränkungen), wobei hier nur potentielle Maßnahmen des Infrastrukturbetreibers analysiert werden können. Ein Austausch von gestörtem rollendem Material wäre z. B. eine Aufgabe des Verkehrsbetriebes und fällt nicht in den Aufgabenbereich eines Dispositionssystems (wobei es natürlich seitens der „Außenwelt“ über solch eine Maßnahme informiert werden muß).

Zur Unterstützung dieser Prozesse benötigt das Dispositionssystem eine sehr genaue Fahrtzeitenrechnung (Trassenprognose) und – damit zusammenhängend – ein mikroskopisches gleis- und signalgenaues Infrastrukturmodell sowie genaue Fahrzeugdaten. Für die

Behandlung von Anschluß- und Umlaufkonflikten ist unter Umständen ein korrespondierendes makroskopisches Infrastrukturmodell erforderlich.

Das Dispositionsmodell muß zusätzlich in der Lage sein, flexible Prioritäten bei der Konfliktlösung zu unterstützen, sowie im Falle von für Prioritäten relevanten Störfällen bestimmte, vorgegebene Elementarereignisse für das Prioritätenvergabe- und -optimierungssystem zu generieren. Eine vereinfachende Darstellung der Architektur eines derartigen Dispositionssystems ist Abb. 6-3 zu entnehmen.

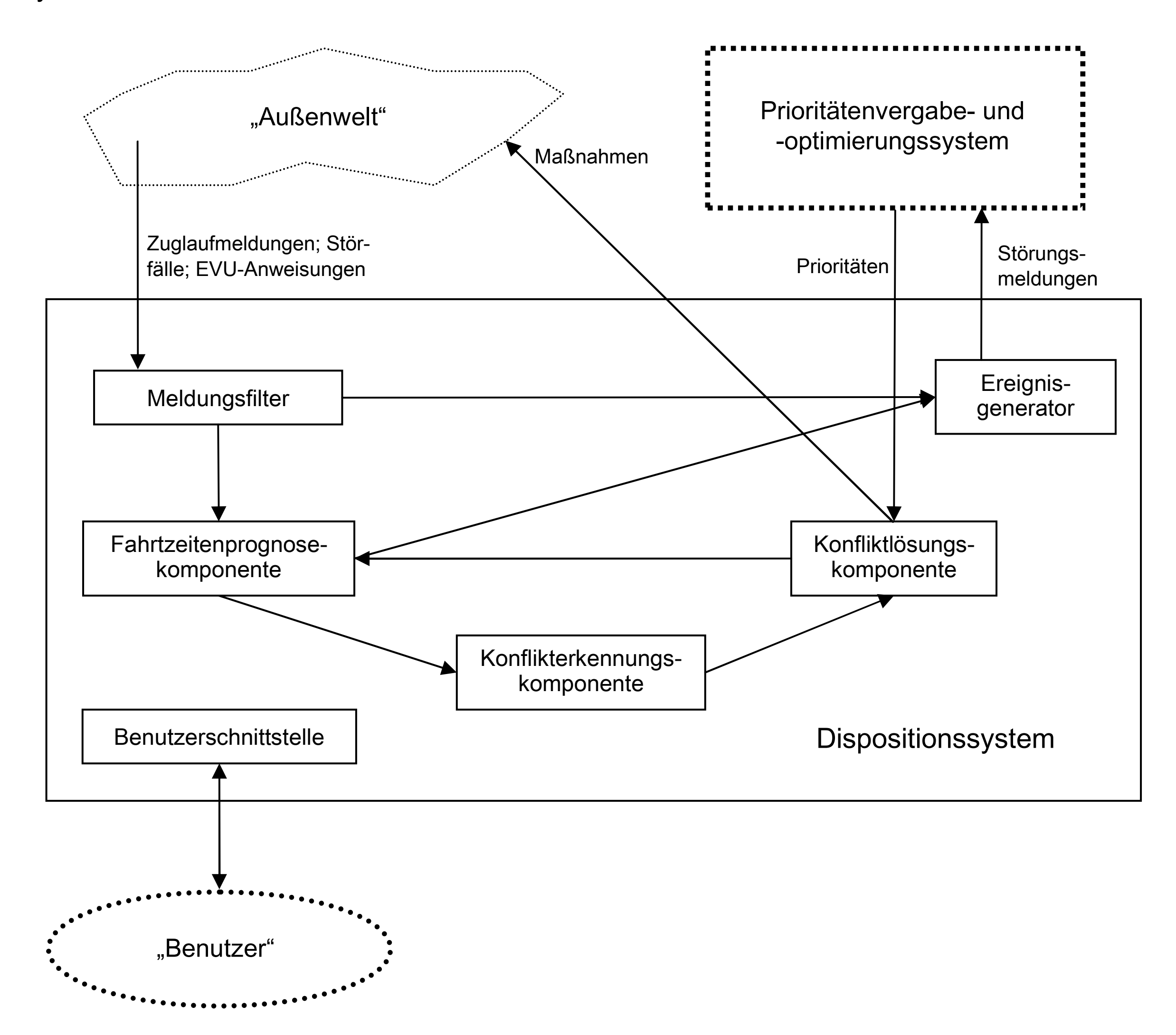


Abb. 6-3: Architektur eines integrierten prioritätenbasierten Dispositionssystems

Das Dispositionssystem besteht somit aus den Teilsystemen Meldungsfilter, Fahrtzeitenprognosekomponente, Konflikterkennungskomponente, Konfliktlösungskomponente und Ereignisgenerator. Zusätzlich ist eine Benutzerschnittstelle vorhanden.

Es existieren weiterhin definierte Schnittstellen zu den externen Systemen Prioritätenvergabe- und -optimierungssystem und der „Außenwelt“.

Der normale Arbeitsablauf des Dispositionssystems ist folgendermaßen: Es enthält genaue Daten über die zugrundeliegende Infrastruktur, die darauf verkehrenden Fahrzeuge, die zugrundeliegenden Trassenwünsche sowie den auf der aktuellen Verkehrslage basierenden

Wunschfahrplan. Über den Meldungsfilter werden anhand der aus der „Außenwelt“ mitgeteilten Zuglaufmeldungen ggf. Abweichungen von diesem Wunschfahrplan ermittelt. Falls so eine Abweichung erkannt wird, muß für den betroffenen Zug eine erneute Fahrtzeitenprognose errechnet werden. Nach Durchführung dieser Prognose erfolgt der wichtige Teilprozeß der Konflikterkennung. Hier wird überprüft, ob eine der drei oben genannten Konfliktarten unter Beteiligung des neu prognostizierten Zuges auftritt. Wenn dies der Fall ist, muß der detektierte Konflikt entsprechend den vom Prioritätenvergabe- und -optimierungssystem übermittelten Prioritäten aufgelöst werden. Dabei muß ggf. eine neue Fahrtzeitenrechnung erfolgen, wodurch neue Konflikte auftreten können. Das genaue Vorgehen in diesem Teilschritt hängt sehr stark von der Art des verwendeten Dispositionssystems ab und soll später detailliert diskutiert werden. Die ermittelten Konfliktlösungen müssen nun wieder in die „Außenwelt“ zur konkreten praktischen Umsetzung zurückgespielt werden.

Neben den klassischen Zuglaufmeldungen müssen noch weitere Informationen aus der „Außenwelt“ berücksichtigt werden: Zum einen Meldungen über temporäre Störungen in Form von eingeschränkten Ressourcen für Fahrweg und Fahrzeug (insbesondere Geschwindigkeitsreduktionen). Diese müssen in geeigneter Weise berücksichtigt werden, so daß für betroffene Züge unter gegebenen Randbedingungen eine neue Fahrtzeitrechnung und darauf aufbauend Konflikterkennung und -lösung erfolgen kann. Zum anderen müssen Vorgaben der Verkehrsbetriebe beachtet werden. Hierbei geht es um speziell fahrzeugseitige Maßnahmen zur Aufrechterhaltung des Verkehrs, wie z. B. Austausch bzw. Außerverkehrziehen eines Fahrzeuges, Einsatz einer Verstärkungsfahrt oder Einrichtung von Schienenersatzverkehren und damit Brechung von Linien. Diese Vorgaben müssen wie manuelle Konfliktlösungen betrachtet werden, die in der Regel auch ein „Anwerfen“ des Konflikterkennungs- und -lösungsprozesses implizieren.

Schließlich ist als „Front-End“ noch eine Benutzerschnittstelle erforderlich, die in erster Linie der anwendungsspezifischen Illustration der Dispositionsfahrpläne dient. Unter Umständen kann hier auch noch ein direkter Einfluß auf die Konfliktlösung ermöglicht werden.

Die dem Dispositionssystem bekannten und für die Festlegung der Prioritäten bedeutsamen Betriebsereignisse, d. h. aus Fahrtzeitenprognosen abgeleitete Planabweichungen sowie Störungsmeldungen, müssen dem Prioritätenvergabe- und -optimierungssystem über eine Schnittstelle mitgeteilt werden. In gleicher Weise werden Änderungen von Prioritäten (aufgrund derartiger Störungen) durch das Prioritätenvergabe- und -optimierungssystem dem Dispositionssystem übermittelt, worauf dieses einen erneuten Konflikterkennungs- und -lösungsprozeß auf Basis der „neuen“ Prioritäten vornehmen muß.

Wie in 2.3 ausführlich dargestellt, unterscheiden sich derzeit verwendete und diskutierte Dispositionssysteme bzgl. der Grundstrategien für die Konfliktlösung. Die Eignung dieser Strategien für die Einbindung des Dispositionssystems in das Gesamtsystem soll im folgenden kurz diskutiert werden.

Der größte Teil der in der Praxis verwendeten Systeme (z. B. die Betriebszentrale der DB Netz AG [Bar01]) arbeitet noch weitgehend auf der Basis **rechnerunterstützter manueller Disposition**. In diesem Falle erfolgen Meldungsfilterung, Fahrtzeitenrechnung und die Kon-

flikterkennung automatisiert, wohingegen die explizite Konfliktlösung weiterhin in der Hand des Bearbeiters bleibt. Bei der Integration eines solchen Dispositionssystems in die Systemumgebung müssen die vom Prioritätenvergabe- und -optimierungssystem vorgeschlagenen Abschnittsweisen Zug- und Anschlußprioritäten dem Benutzer in geeigneter Weise (z. B. farblich) dargestellt werden, so daß dieser einfach eine Konfliktlösung aufgrund dieser Vorgabe entwickeln kann.

Darauf aufbauend finden sich insbesondere im wissenschaftlichen Bereich Dispositionssysteme mit **zentralisierter automatischer Disposition**, welche in 2.3 detailliert dargestellt werden. Hierbei erfolgt zusätzlich auch die Konfliktlösung automatisiert in einer zentralen Einheit. Konfliktlösungsverfahren auf Basis einer Zielfunktion arbeiten nicht prioritätenbasiert und haben daher in diesem Zusammenhang keine Bedeutung. Es verbleiben daher die „regelbasierten“ Verfahren, insbesondere synchrone und asynchrone Simulationen und Expertensysteme. Diese Verfahren sind prinzipiell in der Lage, in die hier entwickelte Systemumgebung integriert zu werden und vom Prioritätenvergabe- und -optimierungssystem übermittelte Prioritäten zu berücksichtigen.

Im Falle von asynchronen Simulationen müssen ggf. für die Behandlung dynamischer Prioritäten Anpassungen vorgenommen werden, da insbesondere die Höherstufung eines Zuges auf seinem Laufweg (z. B. das Einfahren des Güterzuges von einer Personenverkehrsvorrangstrecke auf eine Gütervorrangstrecke) mit klassischen asynchronen Algorithmen nicht geeignet behandelt werden, da der hintere Laufwegsteil vor dem vorderen Teil disponiert werden müßte. Ebenso kann es aus gleichen Gründen Schwierigkeiten bei der Berücksichtigung von Anschlüssen von niederrangigen auf hochrangige Züge geben.

Wegen dieser Problematik sollte der Dispositionshorizont bei den regelmäßig aktualisierten Dispositionsfahrplänen auf einen überschaubaren Zeitraum in der Zukunft beschränkt sein. Wegen der Möglichkeit des jederzeitigen Eintretens weiterer Störungen sind längerfristige Dispositionsfahrpläne ohnehin von nur sehr beschränkter Aussagekraft. Dadurch wird die Zahl der zu einem Zeitpunkt zu beachtenden und für den asynchronen Algorithmus „kritischen“ Höherstufungen und Anschlußaufnahmen deutlich limitiert.

Zur Disposition der Anschlüsse bzw. des hinteren Laufwegteils eines „höhergestuften“ Zuges innerhalb des Untersuchungszeitraums müssen die entsprechenden realen Ankunftszeiten am Grenzbahnhof zwischen den beiden Prioritätenabschnitten zu Beginn des Verfahrens bekannt sein. Dazu können die Prognosen aus dem zuletzt berechneten Dispositionsfahrplan verwendet werden. Sofern keine weiteren Störungen eingetroffen sind, kann von einer weiteren Gültigkeit dieser Ankunftszeiten ausgegangen werden.

Die jeweiligen Laufwegsabschnitte in den einzelnen Prioritätenabschnitten werden entsprechend ihrer Abschnittsweisen Zugprioritäten unter Berücksichtigung der Anschlußprioritäten in den Dispositionsfahrplan während des Untersuchungszeitraums eingelegt, beginnend mit den höchstrangigen Abschnitten. Sollte sich bei Fertigstellung des Dispositionsfahrplans wegen verzögerter Ankunft eines Zuges an einem Übergang zwischen zwei Prioritätenabschnitten ein „Fehler“ bzgl. eines Anschlusses oder nicht mehr möglicher Abfahrtszeit nach einer Höherstufung ergeben, kann die Erstellung bei Bedarf mit den korrigierten prognosti-

zieren Ankunftszeiten wiederholt werden, falls sofort ein gültiger Dispositionsfahrplan benötigt wird.

Letztlich gibt es, wiederum wie bereits in 2.3 ausgeführt, wissenschaftliche Forschungsaktivitäten in Richtung einer **dezentralisierten automatischen Disposition**, die unter Umständen im Zusammenhang mit ohne konkreten Fahrplan betriebenen und/oder selbstorganisierend auf dem Netz verkehrenden Einheiten an Bedeutung gewinnen könnte (vgl. z. B. [Ros92] u. [Rau98]). Hierbei wird die Konflikterkennung und -lösung auf in den Triebfahrzeugen lokale Dispositionsrechner übertragen. Solange das zugehörige Prioritätenvergabe- und –optimierungssystem als zentralisierte Komponente erhalten bleibt, ist auch hier die Einbettung eines derartigen Dispositionssystems in die Gesamtarchitektur denkbar. Die lokalen Rechner würden in diesem Zusammenhang die zentral vergebenen und übermittelten Prioritätenvorgaben in geeigneter Weise abgestimmt mit lokalen Rechnern „benachbarter“ Züge in konkrete konfliktfreie Trassen umsetzen.

Schließlich soll noch explizit darauf hingewiesen werden, daß die Sicherstellung von **Deadlockfreiheit** ein wichtiger Aspekt des Dispositionssystems darstellt. Insbesondere im Zusammenhang mit Strecken, die teilweise eingleisige Streckenabschnitte beinhalten (die in beiden Fahrtrichtungen genutzt werden), kann es bei schlechter Disposition zu Deadlocks kommen (Einfahrt in eingleisigen Abschnitt ohne absehbare Möglichkeit einer Ausfahrt), so daß kein weiterer Verkehr mehr möglich ist bzw. ein unerwünschtes Zurücksetzen einer (oder mehrerer) Einheit erforderlich wird. Diese örtlich eingleisigen Infrastrukturbereiche, in denen Verklemmungen entstehen können, werden infolgedessen als „deadlockgefährdete Bereiche“ bezeichnet. Diese Bereiche können relativ einfach abgegrenzt werden.

Deadlockgefährdete Bereiche müssen vom Dispositionssystem besonders scharf überwacht werden müssen (zügige Konfliktbeseitigung), um zu jedem Zeitpunkt einen durchführbaren und verklemmungsfreien Fahrplan auf Basis der vorhanden ermittelten Prioritäten zu generieren. Im Zweifel hat die Flüssigkeit des Verkehrs (Deadlockfreiheit) selbstverständlich Vorrang vor dem Einhalten der (kommerziell relevanten) Prioritäten.

6.1.3 Anforderungen an das zugehörige Qualitätsüberwachungssystem

Ziel von Bonus-Malus-Regelungen ist es, Verspätungen sowie das Einhalten von vorgegebenen Anschlüssen zahlenmäßig zu gewichten. Details hierzu können 2.2 entnommen werden. Ziel des Infrastrukturbetreibers wird es sein, das Auftreten von Bonus-Malus-Zahlungen zu vermeiden bzw. zu reduzieren. An diesem Interesse muß die Prioritätenvergabe und damit die Disposition ausgerichtet werden. Es sind folgende Bonus-Malus-Regelungen möglich:

- Bonus-Malus-Zahlung relativ zur entstehenden Verspätung an einem vorgegebenen Referenzpunkt
- Bonus-Malus-Zahlung relativ zum Einhalten bzw. Brechen eines bestimmten Anschlusses

Im Falle der Verspätungen kann eine Bonus-Malus-Funktion in Abhängigkeit von den Verspätungsminuten am Referenzpunkt festgelegt werden. Für Anschlüsse werden Absolutwerte für Einhalten bzw. Brechen des Anschlusses bestimmt.

Ein in die hier beschriebene Systemumgebung zu integrierendes Qualitätsüberwachungssystem dient der Bewertung der Resultate des aufgrund der vorgegebenen Prioritäten arbeitenden Dispositionssystems. Die Bewertung erfolgt auf Basis von seitens des Benutzers vorgegebenen Bonus-Malus-Regelungen. Die manuelle Bearbeitung dieser Regelungen soll wiederum mit Hilfe einer Benutzerschnittstelle ermöglicht werden.

Es arbeitet durchgängig im Hintergrund, und seine Kernaufgabe besteht darin, die von der „Außenwelt“ mitgeteilten Verkehrsprotokollierungen begleitend zum und unabhängig vom operativen Geschäft anhand der Bonus-Malus-Regelungen auszuwerten.

Bei Auswertung der Verkehrsprotokolle werden entsprechend der Bonus-Malus-Regelungen für tatsächlich resultierende Bonus-Malus-Zahlungen Ereignismeldungen generiert. Ferner werden die Bonus-Malus-Zahlungen durch das Qualitätsüberwachungssystem auf die eigentlich verursachende Prioritätenabschnitte und Anschlüsse entsprechend der Betriebsprotokolle zugewiesen (vgl. 5.2). Es erfolgt in Prioritätenabschnitten sowohl eine Zuordnung auf die relativ höherrangigen wie auch die relativ niederrangigen Züge. Bei Anschlüssen erfolgt eine Zuordnung auf nicht eingehaltene sowie auf eingehaltene Anschlüsse. Die Bonus-Malus-Zahlungen implizieren einen potentiellen Bedarf zur Prioritätenoptimierung.

Das Qualitätsüberwachungssystem nimmt selber nicht direkt auf die Prioritätenoptimierung Einfluß, welche ausschließlich beim Prioritätenvergabe- und -optimierungssystem verbleibt. Dem Qualitätsüberwachungssystem obliegt es nur, dieses über bestimmte Auffälligkeiten und deren Ursachen zu informieren. Was das Prioritätenvergabe- und -optimierungssystem mit diesen Beobachtungen macht, berührt das Qualitätsüberwachungssystem nicht weiter. Eine vereinfachende Darstellung der Architektur eines derartigen Qualitätsüberwachungssystems ist Abb. 6-4 zu entnehmen.

Das Qualitätsüberwachungssystem besteht somit aus den Komponenten Detektor für Bonus-Malus-Zahlungen aus den Protokolldaten, Ermittler für eigentliche Verursacher von Bonus-Malus-Zahlungen auf verursachende Prioritätenabschnitte und Anschlüsse, sowie dem Bonus-Malus-Ereignisgenerator und einer Benutzerschnittstelle. Es existieren weiterhin definierte Schnittstellen zum Prioritätenvergabe- und -optimierungssystem und der „Außenwelt“. Das Qualitätsüberwachungssystem erhält permanent Protokolldaten (Planabweichungen) aus der „Außenwelt“, die bearbeitet werden müssen. Die generierten Bonus-Malus-Ereignismeldungen werden dem Prioritätenvergabe- und -optimierungssystem mitgeteilt. Das Qualitätsüberwachungssystem arbeitet zeitlich und organisatorisch unabhängig vom Prioritätenvergabe- und -optimierungssystem. Eine Synchronisation der Prozesse ist nicht erforderlich.

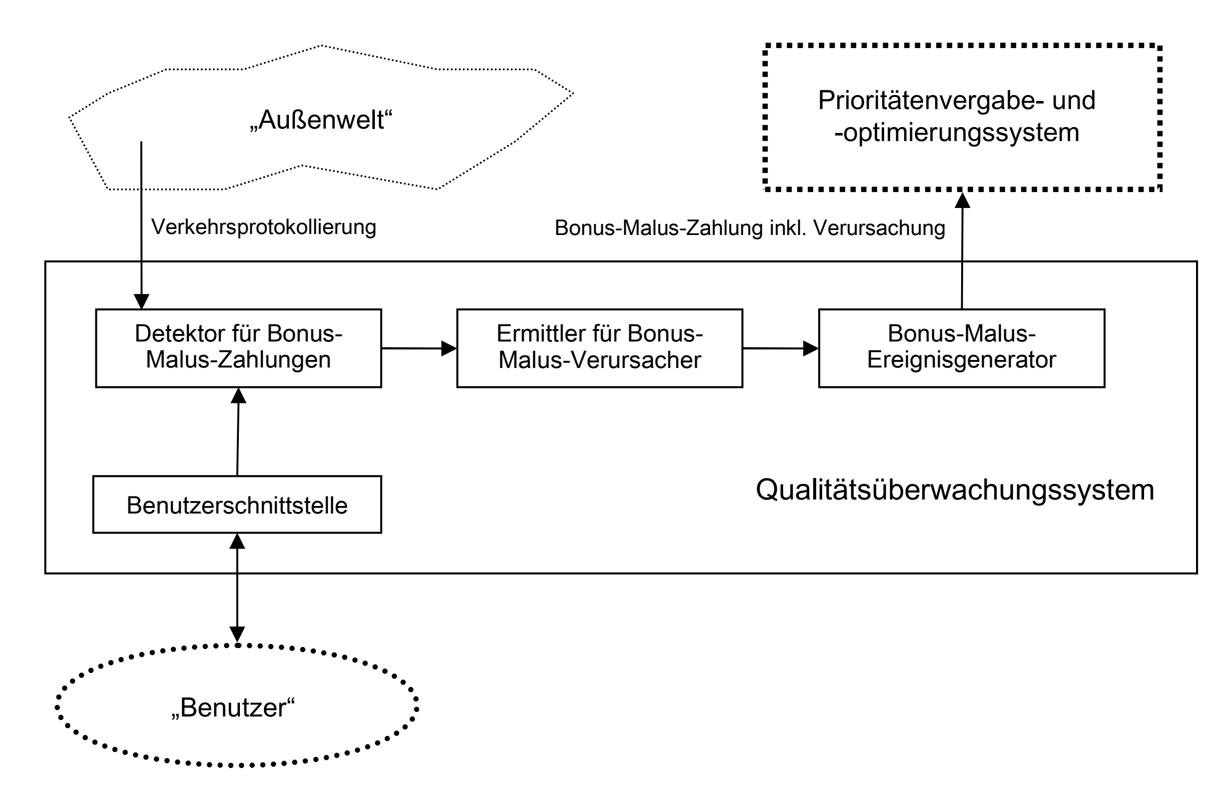


Abb. 6-4: Architektur eines integrierten Qualitätsüberwachungssystems

6.2 Grundzüge eines regelbasierten Ansatzes zur Prioritätenvergabe- und -optimierung

In der soeben vorgestellten Gesamtarchitektur müssen größere Datenmengen (insbesondere über Ereignisse und Zustände) verwaltet werden. Im Prioritätenvergabe- und -optimierungssystem sollen die bei der Disposition zu verwendenden Abschnittsweisen Zug- und Anschlußprioritäten für planmäßigen Verkehr und für Spezielle Betriebszustände jederzeit abrufbar sein und „im Hintergrund“ die anhand des Bonus-Malus-Systems erreichten Ergebnisse langfristig weiter optimiert werden.

Im Rahmen der Prioritätenoptimierung spielen Regeln verschiedener Art eine wichtige Bedeutung. Zum einen existiert eine Vielzahl politischer, ökonomischer oder betrieblicher „Spielregeln“ für die Festlegung von Dispositionsprioritäten. So besteht z. B. das Verbot der Diskriminierung bestimmter Verkehrsunternehmen. Andererseits kann es beispielsweise Vorrangstrecken für bestimmte Verkehrskategorien geben (z. B. den Güterverkehr oder S-Bahn-Verkehr), um eine entsprechende Betriebsqualität sicherzustellen. Derartige Regeln werden nicht einmalig „für die Ewigkeit“ erstellt, sondern können häufigeren Anpassungen durch politische oder betriebliche Entscheidungsträger ausgesetzt sein, die zu entsprechenden Anpassungen des Verfahrens an die geänderten Vorgaben führen.

Zum anderen spielen Ereignisse und die Reaktion auf derartige Ereignisse eine wichtige Rolle im Prioritätenoptimierungsverfahren. In diesem Zusammenhang müssen einerseits wiederum Regeln definiert werden, welche Ereignisse zu möglichen Aktivitäten des Prioritätenvergabe- und -optimierungssystems führen, und anderseits Regeln, die diese Reaktionen

des Systems genauer spezifizieren. Auch hier sind regelmäßige Anpassungen (z. B. im Rahmen der Systemevaluierung) möglich.

Aufgrund der starken Bedeutung dieser verschiedenen Regeltypen für das Anwendungsgebiet wird die Entwicklung eines regelbasierten Prioritätenvergabe- und -optimierungssystems vorgeschlagen. Das System wird dabei unabhängig von den im Einsatzgebiet konkret implementierten Regeln modelliert und konstruiert. Bereits existierende Datenbankmanagementsysteme (DBMS) mit standardisierten Komponenten zur Verwaltung und Bearbeitung derartiger Regeln können in das System integriert werden und ermöglichen damit eine wirtschaftliche Entwicklung und verhältnismäßig einfache langfristige Wartung des gesamten Systems. Die entsprechenden Prioritätenregeln können über eine Schnittstelle dem System übergeben werden und jederzeit (z. B. aufgrund neuer politischer Rahmenbedingungen) geändert werden, ohne daß umfangreiche und kostspielige Anpassungen des eigentlichen Prioritätenvergabe- und -optimierungssystems selber erfolgen müssen.

Die Regeln repräsentieren die spezifischen Details der eigentlichen Optimierungsprozesse im Rahmen der im vorigen Abschnitt aufgezeichneten Grundstruktur und können entsprechend der Ausführungen in 3.3 in drei Kategorien unterteilt werden:

1. **Normative Regeln**

   Normativen Regeln dienen generell der Definition von in-/konsistenten Zuständen einer Datenbank.

   Im Rahmen des Prioritätenvergabe- und -optimierungssystems können normative Regeln zur Definition von für Prioritätenvergabe relevanten festen Rahmenvorschriften (politischen Vorgaben) zur Disposition im jeweiligen Netz verwendet werden (z. B. grundsätzlicher Vorrang für Personenverkehre, allgemeine Mindestwartezeiten für bestimmte Anschlüsse usw.). Dies wird praktisch durch Definition (un-)zulässiger Wertebereiche für einzelne bzw. Gruppen von Abschnittsweisen Zug- und Anschlußprioritäten realisiert.

2. **Deduktive Regeln**

   Deduktive Regeln dienen der Herleitung ableitbaren Fakten aus explizit definierten Basisfakten.

   Im entworfenen Prioritätenvergabe- und -optimierungssystem können deduktive Regeln insbesondere verwendet werden, um aus den vorab optimierten Standard- und Störfallprioritäten die im jeweiligen Fall zu verwendenden Abschnittsweisen Zug- und Anschlußprioritäten abzuleiten. Ferner können deduktive Regeln verwendet werden zur Ableitung der Toleranzgrenzenüberschreitungen für in bestimmten Prioritätenabschnitten oder bei bestimmten Anschlüssen verursachte Bonus-Malus-Zahlungen, auf die das Prioritätenvergabe- und -optimierungssystem mit der eigentlichen Prioritätenoptimierung reagieren soll bzw. zur Bewertung von Prioritätenvorgaben.

3. **Aktive Regeln**

   Aktive Regeln werden für die Definition von automatisch von der Regelverarbeitungskomponente gestarteten Systemreaktionen auf bestimmte von diesen erkannten Ereignissen benötigt.

   Ein Einsatz von aktiven Regeln findet sich im Prioritätenvergabe und -optimierungssystem zum einen bei Einfügungen und Löschung Spezieller Betriebszustände aufgrund eines signalisierten Betriebsereignisses, bei der planmäßigen und störfallbedingten Vergabe der Prioritäten und bei der Durchführung der eigentlichen Prioritätenoptimierung bei Verletzung von Toleranzschwellen bei Bonus-Malus-Zahlungen auf bestimmten Prioritätenabschnitten oder bei Anschlüssen. Hierbei kann durch die aktive Regel eine manuelle Freigabe des Benutzers vor einer endgültigen Änderung eingefordert werden.

6.2.1 Eine regelbasierte Architektur für Prioritätenvergabe- und -optimierung

Durch Verwendung dieser Regelkonzepte ist es möglich, den eigentlichen Implementierungsaufwand durch Verwendung vorgefertigter regelverarbeitender Module eines DBMS erheblich zu reduzieren. Damit ist es lediglich erforderlich, das „Grundgerüst" des Prioritätenvergabe- und -optimierungssystems eigenständig zu programmieren und wesentliche Teile der eigentlichen Algorithmik auf die entsprechenden Komponenten des DBMS auszulagern. Zudem wird die längerfristige Wartung des Systems erheblich vereinfacht. Bei Änderungen der politischen Rahmenbedingungen oder Optimierungsstrategien muß die eigentliche Systemarchitektur nicht oder nur marginal geändert werden, sondern lediglich mit geringem Aufwand einige in der Datenbank in einem wohldefinierten Format vorliegende Regeln in der Datenbankbenutzerschnittstelle entsprechend angepaßt werden. Alles weitere erledigt dann das DBMS.

Um die Regeln entsprechend verwalten und aktualisieren zu können, soll es eine Möglichkeit zur manuellen Bearbeitung der Regeln mittels der Benutzerschnittstelle geben. Diese Benutzerschnittstelle dient ebenso der Information des Benutzers sowie der Einholung von Freigaben für „kritische" Optimierungsschritte. Die um regelverarbeitende Komponenten erweiterte Architektur des Prioritätenvergabe- und -optimierungssystems ist entsprechend schematisch in Abb. 6-5 dargestellt.

Neben den regulären Standardfunktionalitäten des DBMS muß insbesondere die Unterstützung der normativen, deduktiven und aktiven Regelkonzepte (vgl. 3.4) gegeben sein. Dies ist in bei allen größeren serverbasierten kommerziellen Datenbanksystemen (z. B. Oracle DB, DB2 und MS SQL Server) der Fall. Die Komponenten des Datenbanksystems wickeln in diesem Fall neben den genannten Standardfunktionalitäten zentrale algorithmische Aktivitäten des Prioritätenvergabe- und -optimierungssystems ab.

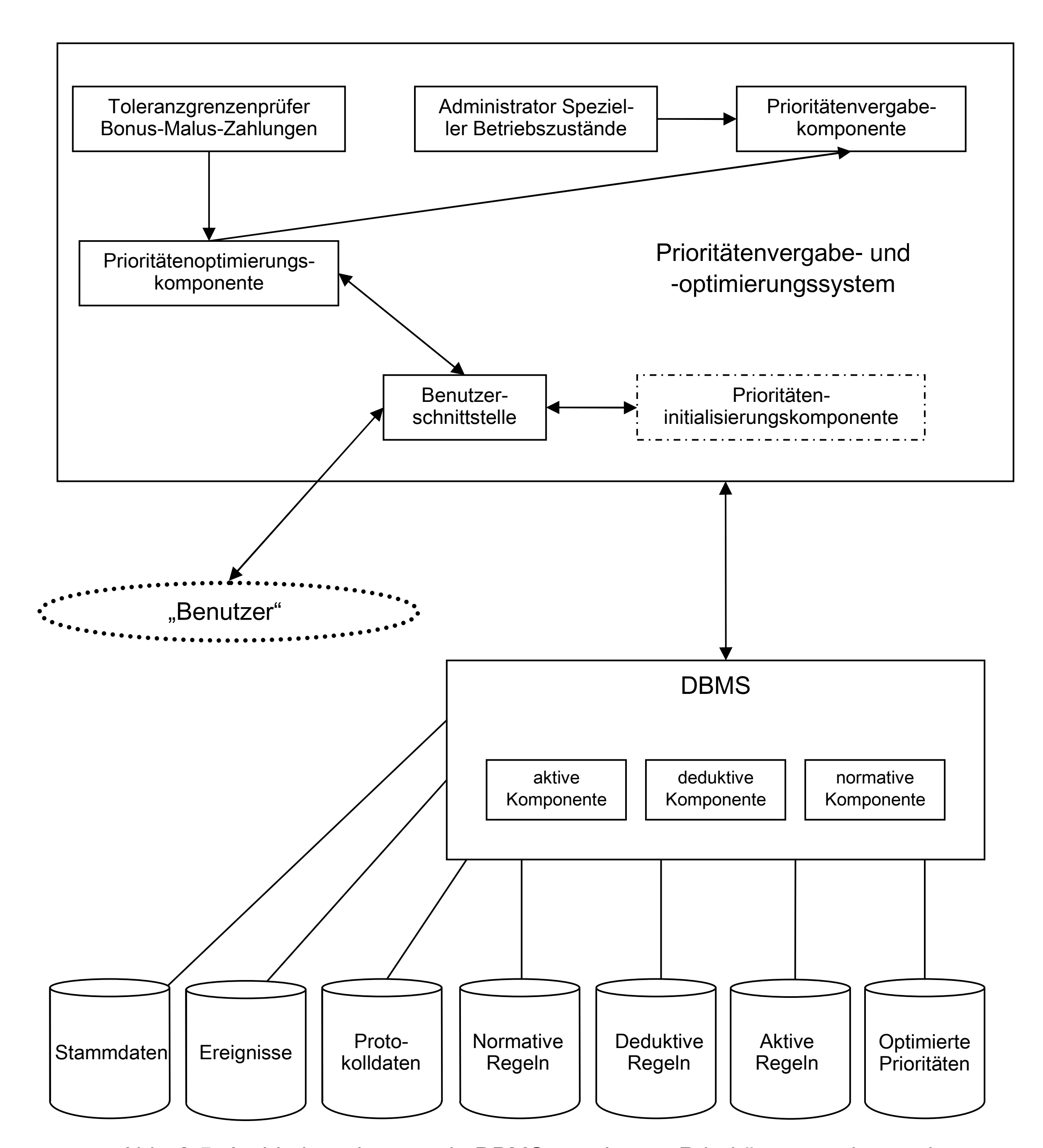


Abb. 6-5: Architektur des um ein DBMS erweiterten Prioritätenvergabe- und -optimierungssystems

Der Aufgabenbereich des DBMS liegt somit bei

- der Verwaltung der Stammdaten der Eisenbahn (makroskopische Infrastruktur, Zugdaten, Fahrplandaten etc.),
- der Verwaltung von durch das Dispositions- bzw. Qualitätsüberwachungssystem übermittelten Betriebsereignisse und Bonus-Malus-Zahlungen inkl. aktuell gültiger Spezieller Betriebszustände,
- die Verwaltung von optimierten Standard- und Störfallprioritäten für Züge und Anschlüsse, die im Bedarfsfalle entsprechend vergeben werden können,

- der längerfristigen Protokollierung von
    - Ereignissen und Speziellen Betriebszuständen,
    - vergebenen Prioritäten und
    - Bewertungen von Prioritätenvergaben,
- der Verwaltung und Überwachung der normativen Regeln zur Definition zulässiger Wertebereiche für Abschnittsweise Zug- und Anschlußprioritäten,
- der Verwaltung deduktiver Regeln und Ableitung der Daten für
    - die Bewertung vergangener Prioritätenvergaben und Feststellung von Toleranzgrenzenüberschreitungen für auf Prioritätenabschnitte bzw. in Anschlüssen verursachte Bonus-Malus-Zahlungen und
    - die Ermittlung der praktisch zu vergebenen Abschnittsweisen Zug- und Anschlußprioritäten aus den vorab optimierten Standard- und Störfallprioritäten und den aktuell gültigen Speziellen Betriebszuständen,
- der Verwaltung der aktiven Regeln und Auslösung relevanter Systemreaktionen für
    - die Durchführung von Prioritätenvergaben bei Anforderung durch das Dispositionssystem bzw. bei Eintritt Spezieller Betriebszustände,
    - die Einfügungen, Aktualisierungen und Löschungen von Speziellen Betriebszuständen und
    - die eigentlichen Prioritätenoptimierung, die bei Überschreitung von Toleranzgrenzen für in Prioritätenabschnitten oder durch Anschlüsse verursachte Bonus-Malus-Zahlungen durchgeführt wird.

Die zentrale Aufgabe beim Entwurf des Prioritätenvergabe- und -optimierungssystems ist es somit, auf geeignete Weise die Optionen des DBMS anzusprechen und auszunutzen.

6.2.2 Schnittstellenerweiterungen

Schließlich sollen noch einzelne offene Punkte zu den Schnittstellen des Prioritätenvergabe- und -optimierungssystems unter dem Gesichtspunkt des Entwurfs mit Hilfe eines regelverarbeitenden DBMS diskutiert werden.

Die **Schnittstelle zum Dispositionssystem** ist entscheidend für das Gesamtkonzept zur Zugdisposition mit Hilfe von optimierten Prioritäten.

Das Dispositionssystem enthält in seiner eigenen Datenbank sehr detaillierte Daten zur Eisenbahninfrastruktur, zu den Zügen, den ursprünglichen Sollfahrplänen und den jeweils gültigen Dispositionsfahrplänen. Ferner speichert es die vom Prioritätenvergabe- und -optimierungssystem übermittelten Prioritäten. Die Schnittstelle dient in der einen Richtung zur Signalisierung von Betriebsereignissen und Prioritätenanfragen an das Prioritätenvergabe- und -optimierungssystem und in der Gegenrichtung zur Rückspielung von zu verwendenden Abschnittsweisen Zug- und Anschlußprioritäten in das Dispositionssystem als Vorgabe für Konfliktlösungen. Dabei ist folgendes bzgl. der verschiedenen Datentransfers zu beachten:

- **Unterschiedliche Infrastrukturmodelle**

  Während im Prioritätenvergabe- und -optimierungssystem die in Kapitel 4 beschriebene makroskopische Infrastrukturmodellierung verwendet wird, benötigt das Dispositionssystem für seine Konflikterkennungs- und -lösungsprozesse ein signal- und weichengenaues mikroskopisches Modell. Um die Kommunikation der Systeme zu ermöglichen, muß eine Konvertierung zwischen diesen beiden Infrastrukturmodellen ermöglicht werden, damit das Prioritätenvergabe- und -optimierungssystem die vom Dispositionssystem signalisierten Betriebsereignisse interpretieren kann. So müssen beispielsweise Signalstörungen für das Prioritätenvergabe- und -optimierungssystem in eine entsprechende Streckenabschnittsstörung transformiert werden. Diese Aufgabe kann vom Dispositionssystem wahrgenommen werden, da dieses das detailliertere Datenformat verwaltet und entsprechend einfach die zugehörigen Objekte des makroskopischen Ansatzes generieren kann (interne Konvertierungsmöglichkeit) und die Ereignisse direkt im „makroskopischen Format“ generieren kann.

  Ferner besteht die Möglichkeit, einen eigenen Konverter zwischen das Dispositions- und Prioritätenvergabe- und -optimierungssystem zwischenzuschalten. Dieser empfängt die Ereignisse eines Systems und wandelt sie in dem anderen System verständliche Ereignisse um. Dementsprechend ist er über die Infrastrukturmodelle beider Systeme informiert und kann entsprechende Transformationen aus dem mikroskopischen in das makroskopische Modell vornehmen.

  Als alternative Architektur bietet sich schließlich eine gemeinsame Infrastrukturdatenbank für Dispositionssystem und Prioritätenvergabe- und -optimierungssystem an. Dabei kann die vom Prioritätenvergabe- und -optimierungssystem verwendete makroskopische Sicht ggf. mittels deduktiver Regeln aus dem vom Dispositionssystem verwendeten mikroskopischen Modell abgeleitet werden. Beide Systeme können lesend auf dieselbe Datenbasis zugreifen. Es entsteht weniger Wartungsaufwand und die Fehleranfälligkeit wegen inkonsistenter Daten ist geringer, so daß dieses Vorgehen bevorzugt werden sollte. Die Ereignissignalisierung kann entsprechend im makroskopischen Format erfolgen.

  Eine detaillierte Ausführung ist jeweils vom tatsächlich verwendeten Dispositionssystem und seiner Infrastrukturmodellierung abhängig und soll hier nicht weiter behandelt werden.

- **Unterschiedliche Granularität der Zug- und Fahrplandaten**

  Hier tritt ein analoges Problem zur Infrastrukturmodellierung auf. Das Dispositionssystem hat aufgrund seiner Aufgaben ein erheblich genaueres Datenmodell zu den Zügen und Fahrplänen als dies für das Datenmodell des Prioritätenvergabe- und -optimierungssystem erforderlich ist. Allerdings kann letzteres genauere Informationen über für das Dispositionssystem nicht so entscheidende verkehrswirtschaftliche Bewertungen der Zugfahrten und Anschlüsse enthalten. Eine Konvertierung zwischen den Modellen läßt sich allerdings durch „Filterung“ der jeweils relevanten Informationen relativ leicht herzustellen, so daß eine Verständigung der Systeme bzgl. der übermittelten Betriebsereignisse und Prioritäten ermöglicht werden kann. Allerdings muß die Äquivalenz der Stammdaten

in beiden Systemen (insbesondere bei der Einspielung von Fahrplanänderungen) gewährleistet sein. Dazu bestehen wiederum die im Zusammenhang mit den Infrastrukturmodellen skizzierten Möglichkeiten.

Sollte eine Architektur mit einer teilweise gemeinsamen Datenbank gewählt werden, besteht ggf. die Möglichkeit, auf eine explizite Übermittlung der vergebenen Abschnittsweisen Zug- und Anschlußprioritäten durch das Prioritätenvergabe- und -optimierungssystem aufgrund einer Anfrage des Dispositionssystems zu verzichten und statt dessen einen direkten aktiven Lesezugriff durch das Dispositionssystem auf die vom Prioritätenvergabe- und -optimierungssystem in der Datenbank abgelegten Daten durchzuführen. Dies führt zu einer Beschleunigung des Vergabeverfahrens.

Die **Schnittstelle mit dem Qualitätsüberwachungssystem** wird in einem geringeren Umfang genutzt. Es dient ausschließlich der Übermittlung von vom Qualitätsüberwachungssystem im realen Verkehr detektierten Bonus-Malus-Zahlungen. Neben dem eigentlichen Betrag und dem Grund der Bonus-Malus-Zahlung, einer Planabweichung bzgl. eines Verkehrshaltes oder eines Anschlusses enthält das Bonus-Malus-Ereignis Angaben über die verursachenden Prioritätenabschnitte (niederrangige und höherrangige Züge) bzw. Anschlüsse (eingehaltene und gebrochene) dieser Zahlung.

Das Qualitätsüberwachungssystem verwaltet in diesem Zusammenhang eine große Menge von aus der „Außenwelt“ übermittelten Verkehrsmeldungen. Ferner enthält es eine ausführliche Aufstellung der Bonus-Malus-Regelungen. Diese werden für Planabweichungen von Ankunfts- und Abfahrtszeiten als vom Abweichungsbetrag abhängige Funktionen definiert. Die Funktionen müssen nicht stetig sein. So besteht z. B. die Möglichkeit, Verspätungen unter 30 Minuten ohne Bonus-Malus-Zahlung zu belegen und darüber hinaus einen konstanten Betrag zu deklarieren. Für jeden Anschluß wird eine boolesche Funktion definiert, abhängig davon ob der Anschluß eingehalten wurde oder nicht.

Grundsätzlich kann das Qualitätsüberwachungssystem auf denselben Stammdaten (Infrastruktur, Züge und Fahrplan) wie das Prioritätenvergabe- und -optimierungssystem arbeiten. Die Anforderungen an die Granularität der Daten sind in beiden Systemen ähnlich. Dementsprechend bietet sich für diesen Zweck der gemeinsame Zugriff auf dieselbe Datenbasis an. Auf diese Weise kann die Kommunikation der Systeme problemlos erfolgen.

Die **Benutzerschnittstelle** dient im Zusammenhang mit der DBMS-basierten Architektur des Prioritätenvergabe- und -optimierungssystems zunächst einmal dazu, ggf. durch aktive Regeln zur Prioritätenoptimierung erforderliche manuelle Freigaben für Manipulationen von Standard- und Störfallprioritäten entgegenzunehmen. Falls eine aktive Regel mit Freigabe-Erfordernis feuert, wird der Benutzer hierüber geeignet informiert und kann der vorgeschlagenen Prioritätenmanipulation dann entweder zustimmen oder diese ablehnen. Im letzteren Fall bleiben die bisher gültigen Standard- und Störfallprioritäten in Kraft.

Eine weitere Aufgabe der Benutzerschnittstelle ist es, im Rahmen der längerfristigen Wartung des Systems die manuelle Pflege der Stammdaten sowie der normativen, deduktiven und aktiven Regeln zu übernehmen. Ziel einer regelbasierten Architektur war es, durch Ein-

bindung von Standard-DBMS-Komponenten den Implementierungsaufwand des Prioritätenvergabe- und -optimierungssystems gering zu halten und den langfristigen Wartungsaufwand zu senken. Wesentliche algorithmische Anpassungen erfolgen somit durch Bearbeitungen der Regelbasis. Die manuellen Bearbeitungen der Regeln sollte möglichst einfach ermöglicht werden, ggf. mit Unterstützung graphischer Elemente.

6.3 Spezifikation der Prozeßabläufe

Gemäß den Ausführungen der vorausgehenden Abschnitte ist das entwickelte Prioritätenvergabe- und -optimierungssystem in der Regel „passiv“, d. h. es wartet auf äußere Reize, aufgrund derer es zu arbeiten beginnt. Diese Aktivierung des Systems erfolgt entweder durch das angeschlossene Dispositionssystem oder das angeschlossene Qualitätsüberwachungssystem zur Bewertung des Verkehrs anhand eines vorgegebenen Bonus-Malus-Systems. In Sonderfällen kann auch durch Interaktion des Benutzers mit dem System eine derartige Aktivierung erfolgen. Die eigentlichen Kernkompetenzen des Prioritätenvergabe- und -aktivierungssystems sind die Verwaltung, Bereitstellung und Optimierung von Abschnittsweisen Zug- und Anschlußprioritäten für das Dispositionssystem auf der Basis von betrieblichen Auswertungen des Qualitätsüberwachungssystems.

Dabei kann man die zentrale Funktionalität des Systems prinzipiell in zwei zentrale (voneinander weitgehend unabhängige) Prozesse unterteilen:

A. Echtzeitvergabe der Abschnittsweisen Zug- und Anschlußprioritäten durch Übermittlung von intern vorab optimierten und verwalteten Standard- und Störfallprioritäten an das Dispositionssystem, unterschieden in

    A.1 Planmäßige Prioritätenvergabe und

    A.2 Verwaltung der Speziellen Betriebszustände und störfallbedingte Prioritätenvergabe

B. Längerfristige Optimierung der intern verwalteten Standard- und Störfallprioritäten auf Basis des vorgegebenen Bonus-Malus-Systems, unterschieden in

    B.1 (Partielle) Prioritäteninitialisierung und

    B.2 Prioritätenoptimierung

Während die „A-Prozesse“ in erster Line in Zusammenarbeit mit dem Dispositionssystem ablaufen, werden die „B-Prozesse“ in Kooperation vom Qualitätsüberwachungssystem bzw. durch Benutzerinteraktion ausgelöst. Die endgültige operative Aktivierung der hier optimierten Prioritäten erfolgt allerdings erst zu vorab festgelegten Optimierungszeitpunkten, um eine längerfristige Bewertung der Resultate entsprechend zu ermöglichen.

Initialisiert werden alle Prozesse durch externe Ereignissignalisierungen an das Prioritätenvergabe- und -optimierungssystem, die in der Regel vom Dispositions- bzw. Qualitätsüberwachungssystem erkannt und signalisiert werden. Eine Besonderheit stellt der Prozeß B.1 dar; dieser wird im Rahmen der Initialisierung des gesamten Systems für alle Prioritäten

aufgerufen. Die weitere Optimierung der erstmalig initialisierten Prioritäten erfolgt mit im fortlaufenden Betrieb mit Prozeß B.2. Der Prozeß B.1 wird im weiteren Systembetrieb im Zusammenhang mit partiellen erneuten Initialisierungen im Fall von umfangreicheren Änderungen von Eingangsdaten benötigt.

Das Prioritätenvergabe- und -optimierungssystem hat die Aufgabe, auf derartige Ereignisse zu warten und entsprechend die oben genannten Prozesse durchzuführen. Dabei muß sichergestellt werden, daß kein signalisiertes Ereignis verlorengeht (Einrichtung eines Zwischenpuffers). In Abb. 6-6 ist dies in einem generellen dynamischen Modell illustriert:

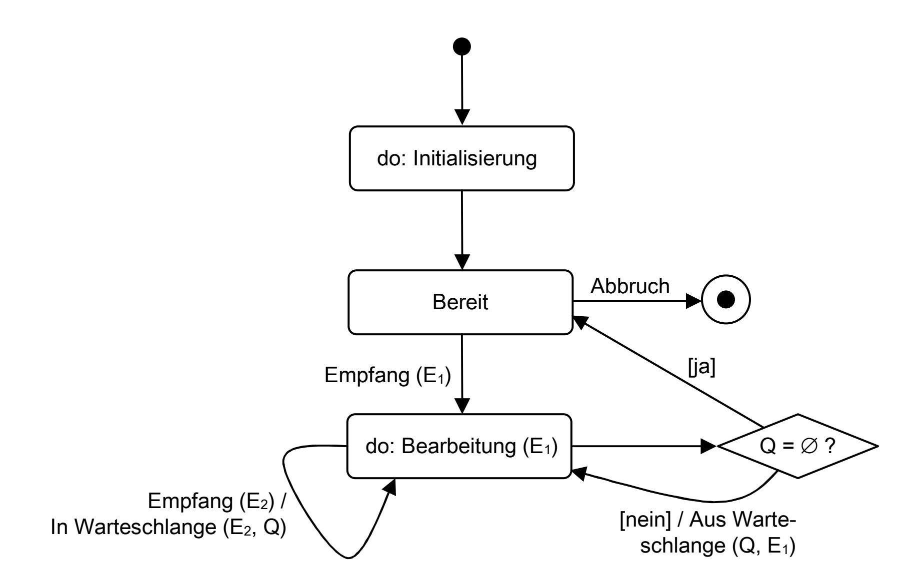


Abb. 6-6: Allgemeines dynamisches Modell des Prioritätenvergabe- und -optimierungssystems

Im folgenden sollen nun die Systeminitialisierung und die Ereignisbearbeitung genauer spezifiziert werden.

6.3.1 Initialisierung

Im Rahmen der Initialisierung des Prioritätenvergabe- und -optimierungssystems müssen die benötigten Stammdaten eingelesen sowie die Prioritäten als Ausgangsbasis für die spätere Optimierung initialisiert werden. Das Einlesen der Stammdaten ist stark vom verwendeten Dispositions- und Qualitätsüberwachungssystem bzgl. dessen Architektur bzgl. Datenbankzugriffen abhängig und wird daher hier nicht detailliert beschreiben. Wichtig dabei ist, die Stammdaten des Prioritätenvergabe- und -optimierungssystem geeignet mit den Daten des Dispositions- bzw. Qualitätsüberwachungssystem zu „verknüpfen“, um eine spätere Kommunikation sicherstellen zu können.

Bei der Initialisierung der Prioritäten handelt es sich um den oben genannten Prozeß B.1. Um Prioritäten optimieren zu können, muß zunächst einmal eine Grundlage bereitgestellt werden, die entsprechend weiter optimiert werden kann. Es sollen plausible Prioritäten festgelegt werden, die bereits eine gewisse Nähe zu den im weiteren Verlauf gemäß Bonus-Malus-System optimalen Werten haben sollten.

Grundsätzlich sind mehrere Strategien zur Festlegung initialer Prioritäten denkbar. Da der Fokus in dieser Arbeit auf den eigentlichen Optimierungsprozessen liegen soll, wird hier nur ein relativ allgemeines Konzept zur Prioritäteninitialisierung vorgeschlagen und entsprechend auf die weitere Optimierung verwiesen. Vor einem praktischen Einsatz kann auch auf Betriebssimulationen zur weiteren Eichung der Prioritäten mit Hilfe des Optimierungsverfahrens zurückgegriffen werden, bis eine zufriedenstellende Qualität im Rahmen der Simulationen erreicht ist.

Entsprechend wird vorgeschlagen, bei der Initialisierung der Prioritäten zunächst lediglich die Standardpriorität zu ermitteln. Störfallprioritäten werden dementsprechend zunächst nicht definiert, d. h. in Störfällen greifen zunächst auch die Standardprioritäten. (Ausnahmen können ggf. für schwere technische Zugstörungen oder sehr hohe Verspätungen gemacht werden, auf die generell mit niedriger Priorität reagiert werden kann.)

Ferner wird vorgeschlagen, die Abschnittsweisen Zugprioritäten bei Initialisierung zunächst einheitlich für den gesamten Zuglauf vorzugeben. Eine Feinjustierung auf verschiedenen Streckenabschnitten kann später im Rahmen der Optimierung erfolgen. Die Ermittlung der Prioritäten erfolgt folgendermaßen: Jeder Zuglauf wird für sich untersucht. Um einen Referenzwert zu erhalten, werden für alle Halte, in denen für Planabweichungen bei der Ankunftszeit Bonus-Malus-Regelungen definiert sind, die Beträge der Bonus-Malus-Zahlungen für verschiedene festzulegende Verspätungen (wie beispielsweise 15 Minuten, 30 Minuten, 60 Minuten und 120 Minuten) gemittelt. Dadurch ergibt sich für den Zug ein Referenzwert für die Gewichtung der Planabweichung eines Verkehrshaltes.

Der Referenzwert für die Gewichtung eines kompletten Zuglaufes ergibt sich aus den mittleren Referenzwerten für die (bewerteten) Verkehrshalte eines Zuges. Unbewertete Verkehrshalte sollen bzgl. der Berechnung ignoriert werden. Züge, für die keine Bonus-Malus-Regelungen vorliegen, werden gesondert behandelt (s. u.)

Nachdem für alle Züge ein mittlerer Referenzwert über den gesamten Laufweg gebildet wurde, können die Züge nun entsprechend dieses Wertes sortiert werden. Die Züge mit dem höchsten Referenzwert erhalten durchgängig eine Abschnittsweise Zugspriorität von 2 (die höchste Prioritätenstufe bleibt standardmäßig für dringliche Hilfsfahrten o. ä. frei). Die Züge mit dem niedrigsten Referenzwert erhalten die zweitniedrigste Priorität n – 1 (n = Anzahl der Prioritätenstufen). Für alle dazwischen liegenden Werte findet eine lineare Interpolation statt, so daß gilt

$$\Pr(z) = \mathrm{round}((n-3) \cdot \frac{\max(\mathrm{ref}) - \mathrm{ref}(z)}{\max(\mathrm{ref}) - \min(\mathrm{ref})} + 2)$$

Züge, für die kein Referenzwert aufgestellt werden konnte, werden mit der Priorität n (niedrigste Priorität) behandelt. Es wird empfohlen, die Anzahl der Prioritäten (n) nicht zu groß zu wählen, z. B. auf 10 festzulegen. Zu viele Abstufungen sind zum einen nach außen schwerer nachvollziehbar, zum anderen erlauben sie dem Optimierungssystem zu viele Freiheitsgrade für eine mögliche Lösung, was nicht unbedingt zielführend ist.

Für Anschlußprioritäten (maximale Wartezeiten) können relativ einfach Standardwerte festgelegt werden. Falls keine Bonus-Malus-Regelung für den Anschluß vorliegt, wird der Wert der Priorität auf 0 Minuten gesetzt (kein Warten erforderlich). Falls ein Wert festgelegt ist, kann die maximale Wartezeit standardmäßig zunächst im Personenverkehr auf 5 Minuten und im Güterverkehr auf 30 Minuten festgelegt werden. Falls ein sehr hoher Bonus-Malus-Wert für den Anschluß vorgegeben ist (z. B. Nachtzüge o. ä.), kann die maximale Wartezeit auch entsprechend verlängert werden.

Für später neu einzufügende Sonderzüge und begrenzte Fahrplanänderungen bzw. Änderungen des Bonus-Malus-Systems bzw. der normativen Randbedingungen wird das Verfahren der Prioritäteninitialisierung sinngemäß ausschließlich für die betroffenen Züge angewendet. Die lokale Feinjustierung der Abschnittsweisen Zug- und Anschlußprioritäten und die besondere Berücksichtigung von Störfallprioritäten erfolgt dann im Rahmen der späteren regelbasierten Optimierung unter Berücksichtung der tatsächlichen betrieblichen Verhältnisse.

Selbstverständlich muß die Einhaltung der normativen Randbedingungen (politische Einschränkungen des Prioritätenbereichs) vom System beachtet werden. In diesen Fällen kann es auch unter Umständen erforderlich sein, die Prioritäten zu korrigieren oder bereits im Rahmen der Initialisierung Störfallprioritäten zu definieren.

6.3.2 Ereignisbearbeitung

Innerhalb des „Handlungsrahmens“ des Prioritätenvergabe- und -optimierungssystems steht die Reaktion auf einzelne Ereignisse als Schwerpunkt dar. Dabei werden je nach vom Dispositions-, Qualitätsüberwachungssystem oder durch den Benutzer über die Benutzerschnittstelle übermitteltem Ereignis die bereits genannten Prozesse des Systems durchgeführt (vgl. Abb. 6-7).

Zunächst muß der Typ des bearbeiteten Ereignisses bestimmt werden. Bei einer Anfrage bzw. Anforderung von Prioritäten erfolgt eine planmäßige Prioritätenvergabe (A.1), bei einem Betriebsereignis muß ggf. ein Spezieller Betriebszustand einfügt, geändert oder gelöscht werden, und es kann ggf. eine störfallbedingte Prioritätenvergabe erfolgen (A.2). Bei einer Änderung von Eingangsdaten durch den Benutzer (z. B. neuer Zug, Fahrplanänderung, Änderung von normativen Rahmenbedingungen oder Bonus-Malus-Regelungen) erfolgt eine partielle Prioritäteninitialisierung (B.1) und schließlich bei einem Bonus-Malus-Ereignis – zeitverzögert – die Prioritätenoptimierung (B.2).

Im folgenden sollen nun die einzelnen Prozesse genauer dargestellt werden.

Die **planmäßige Prioritätenvergabe** erfolgt auf Anfrage bzw. Anforderung des Dispositionssystems. Ziel ist es, diesem vorab optimierte Abschnittsweise Zug- und Anschlußprioritäten zur Disposition zur Verfügung zu stellen, anhand derer die Lösung von Konflikten durchgeführt werden kann.

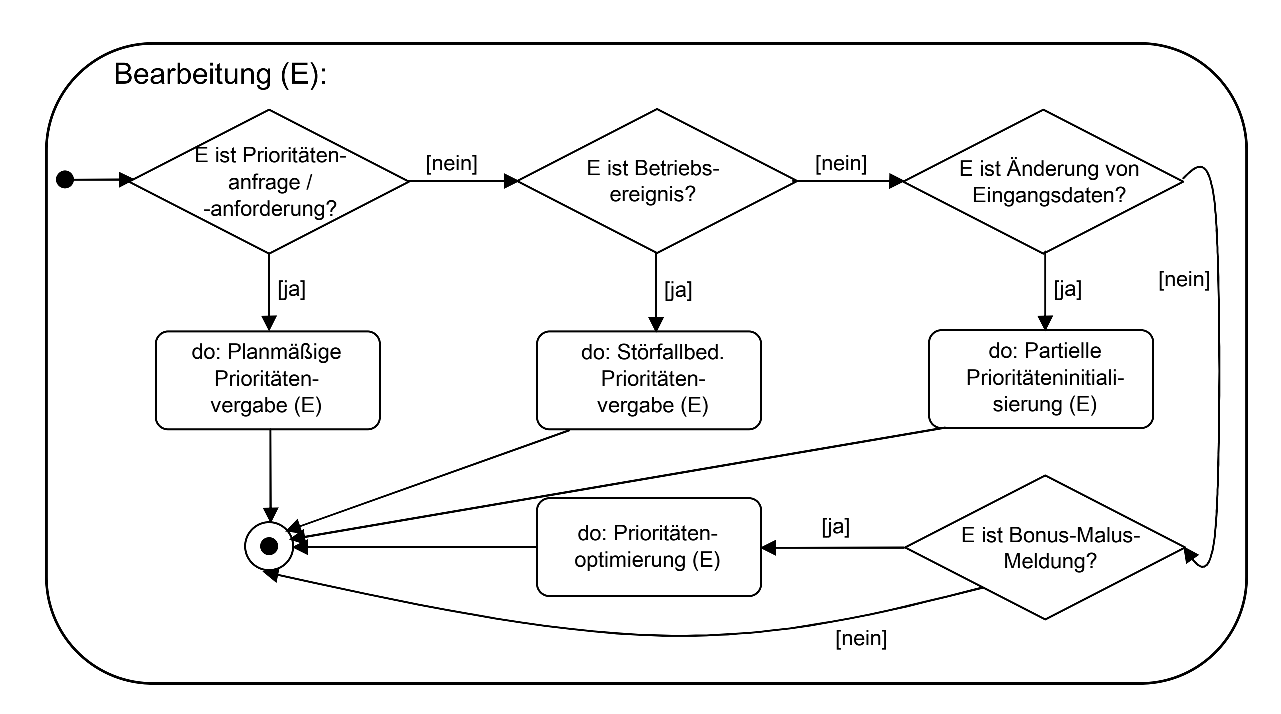


Abb. 6-7: Dynamisches Modell der Ereignisbearbeitung

Dies kann je nach Architektur praktisch zum einen über einen gemeinsamen Zugriff des Prioritätenvergabe- und -optimierungssystems (schreibend) und des Dispositionssystems (lesend) auf eine Prioritätendatenbank realisiert werden. In diesem Falle ist es Aufgabe des Prioritätenvergabe- und -optimierungssystems, in allen Fällen fortwährend konsistente Abschnittsweise Zug- und Anschlußprioritäten für alle Züge und Prioritätenabschnitte sowie alle Anschlüsse zum potentiellen Auslesen durch das Dispositionssystem zu unterhalten. In diesem Falle erfolgt die planmäßige Prioritätenvergabe in einem ereignisungebundenen Hintergrundprozeß des Prioritätenvergabe- und -optimierungssystems. Sämtliche ggf. benötigten Abschnittsweisen Zug- und Anschlußprioritäten müssen rechtzeitig berechnet und die Abrufbarkeit dieser Daten in der Prioritätendatenbank durch das Dispositionssystem bei Bedarf sichergestellt werden.

Alternativ ist es möglich, die Prioritätenvergabe ereignisorientiert mittels Botschaften abzuwickeln. Dabei kann das Dispositionssystem eine vorläufige Anfrage zu Informationszwecken oder eine definitive Prioritätenanforderung für eine Abschnittsweise Zug- oder Anschlußpriorität an das Prioritätenvergabe- und -optimierungssystems übermitteln. Die Anforderung durch das Dispositionssystem erfolgt dabei unmittelbar vor Einfahrt des betroffenen Zuges in den Prioritätenabschnitt (für Abschnittsweise Zugprioritäten) sowie bei Ankunft des Zuges in einen Bahnhof mit Anschlußverbindungen (für Anschlußprioritäten). Nach der endgültigen Vergabe ist die Priorität dann normalerweise festgeschrieben und kann nur noch beim Eintreten schwerer Störfälle geändert werden.

Der Prozeß der Vergabe erfolgt folgendermaßen (vgl. Abb. 6-8): Nach Eingang der Anfrage bzw. Anforderung an das Prioritätenvergabe- und -optimierungssystem (bei einer Implementierung mit Botschaften) bzw. bei Bedarf (bei einer Realisierung mit einem Hintergrundprozeß) wird überprüft, ob derzeit bzgl. der zu übermittelnden Priorität ein Spezieller Betriebs-

zustand im System registriert ist. Wenn das nicht der Fall ist, wird die Standardpriorität für die Disposition übermittelt. Falls doch, muß die für entsprechend dem Gewicht des Typs bedeutendsten Spezieller Betriebszustand (wenn mehrere gleichzeitig vorhanden sind – Sicherstellung der Eindeutigkeit von Störfallprioritäten) definierte Störfallpriorität dem Dispositionssystem mitgeteilt bzw. bereitgestellt werden. Die praktische Implementierung kann mit Hilfe deduktiver und aktiver Regeln erfolgen (vgl. 6.4.2 und 6.4.5).

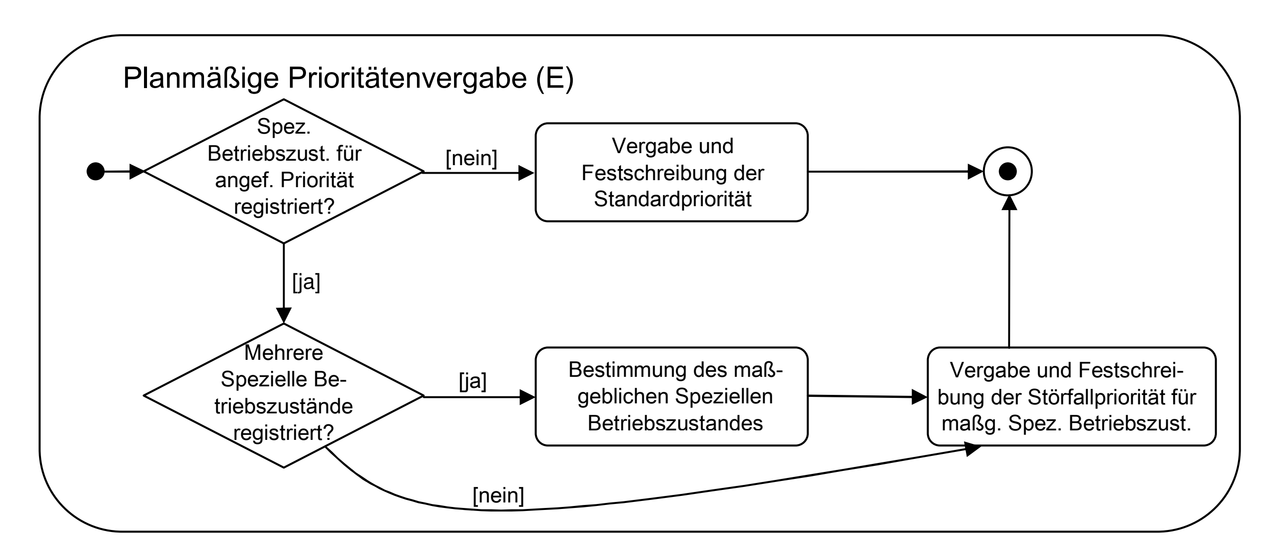


Abb. 6-8: Dynamisches Modell der planmäßigen Prioritätenvergabe

Die **störfallbedingte Prioritätenvergabe** erfolgt bei Signalisierung der Initialisierung, Aktualisierung oder Terminierung eines Speziellen Betriebszustandes mit hohem Gewicht des Typs durch das Dispositionssystem. Ziel ist die Ad-hoc-Reaktion auf plötzlich eintreffende bzw. sich ändernde oder endende, für die Prioritätenvergabe relevante Spezielle Betriebszustände. Unter Umständen müssen dem Dispositionssystem für eigentlich bereits festgeschriebene Prioritäten aufgrund wesentlich geänderter Betriebsbedingungen kurzfristig Änderungen mitgeteilt werden. (Dies betrifft insbesondere Abschnittsweise Zugprioritäten; Anschlußprioritäten sind praktisch nur bei stark eingeschränkten Ressourcen des aufnehmenden Zuges betroffen.)

Diese Prioritätenänderungen „unter rollendem Rad“ aufgrund von kurzfristig eingetroffenen bzw. beendeten Störfällen erfolgen durch Generierung einer entsprechenden Botschaft an das Dispositionssystem, das darauf entsprechend durch Ad-hoc-Umplanung des Dispositionsfahrplans (soweit dies praktisch noch umsetzbar ist) reagieren muß. Eine bei planmäßiger Prioritätenvergabe mögliche Realisierung mittels eines Datenbanklesezugriffs des Dispositionssystems macht hier keinen Sinn, da sehr kurzfristige Reaktionszeiten des Dispositionssystems erforderlich sind. Eine störfallbedingte Prioritätenvergabe erfolgt nur für bereits angeforderte und festgeschriebene Prioritäten, lediglich vorab angefragte Prioritäten werden nicht korrigiert. Hier werden erst bei der nächsten Anfrage bzw. der endgültigen Anforderung die entsprechenden Störfallprioritäten zur Disposition übermittelt.

Der Prozeß der störfallbedingten Vergabe erfolgt dann im Detail folgendermaßen (vgl. Abb. 6-9): Bei Meldung eines Betriebsereignisses durch das Dispositionssystem soll dieses zunächst einmal genauer untersucht werden. Wenn das Ereignis einen Speziellen Betriebs-

zustand initialisiert, muß dieses durch das Prioritätenvergabe- und -optimierungssystem intern als aktiver Spezieller Betriebszustand mit dem entsprechenden Gewicht des Typs registriert werden. Man beachte in diesem Zusammenhang, daß Terminierungsereignisse für punktuelle Störungen bzgl. Spezieller Betriebszustände als initialisierend betrachtet werden und ebenso wie vergangene Planabweichungen nicht mehr aktualisiert oder terminiert werden können.

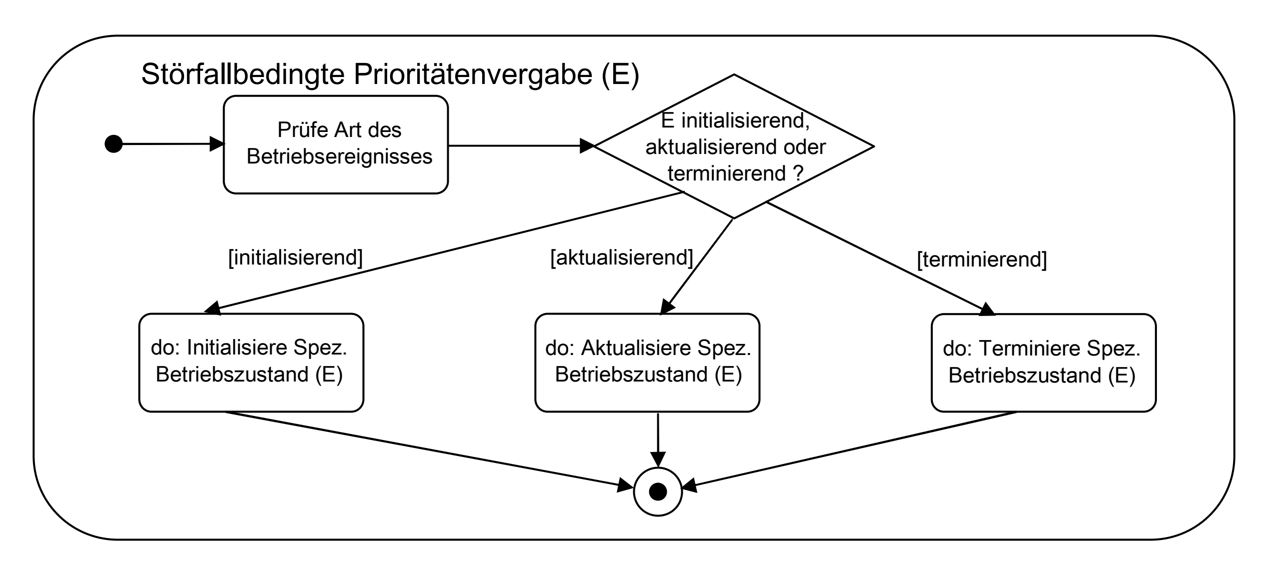


Abb. 6-9: Dynamisches Modell der störfallbedingten Prioritätenvergabe

Wenn das Betriebsereignis einen Speziellen Betriebszustand aktualisiert, muß der bereits registrierte Zustand entsprechend angepaßt werden. Nach der Registrierung bzw. Aktualisierung der Speziellen Betriebszustände muß für davon potentiell betroffene angeforderte und festgeschriebene Abschnittsweisen Zug- und Anschlußprioritäten überprüft werden, ob eine nachträgliche Anpassung auf eine entsprechende Störfallpriorität erforderlich ist.

Bezieht sich der registrierte oder aktualisierte Spezielle Betriebszustand auf einen bestimmten Zug, so betrifft dies die aktuelle Abschnittsweise Zug- und Anschlußpriorität dieses Zuges sowie alle Anschlußprioritäten von Anschlüssen, bei denen der gestörte Zug Aufnehmerfunktion wahrnimmt. Ist der Spezielle Systemzustand streckenbezogen (Streckenstörung), betrifft dies alle Abschnittsweisen Zugprioritäten im betroffenen Prioritätenabschnitt und alle Anschlußprioritäten von Anschlüssen in Stationen, die an der Grenze des betroffenen Prioritätenabschnittes liegen und in denen der aufnehmende Zug in den gestörten Abschnitt einfährt.

Die potentiell betroffenen Prioritäten werden einzeln untersucht (vgl. Abb. 6-10). Es wird geprüft, ob unabhängig vom gerade initialisierten/aktualisierten Zustand für die betroffene Priorität weitere relevante Spezielle Betriebszustände gültig sind. Wenn dies der Fall ist, ist der Spezielle Betriebszustand mit dem höchsten Gewicht des Typs maßgeblich, für den eine Störfallpriorität definiert wurde. Dazu wird überprüft, ob für die betroffene Abschnittsweise Zug- oder Anschlußpriorität bereits eine andere (davon unabhängige) Störfallpriorität für einen gleichzeitigen Zustand mit höhergewichtigen Speziellen Betriebszustandstyp für die Disposition festgeschrieben wurde. Damit soll die Eindeutigkeit der Prioritäten bei sich überlagernden Speziellen Betriebszuständen sichergestellt werden. Falls ja, kann die weitere

Prüfung entfallen. Ansonsten wird, wenn der Spezielle Betriebszustandstyp ein ausreichend hohes Gewicht für eine kurzfristige Prioritätenänderung aufweist, die neue Störfallpriorität vorgemerkt.

Bei störfallbedingten Prioritätenvergaben „unter rollendem Rad" soll über die Benutzerschnittstelle eine manuelle Freigabe angefordert werden, bevor diese kurzfristige und unter Umständen schwierig zu implementierende Änderung festgeschrieben wird. Wenn diese Freigabe vom Benutzer gegeben wird, wird die Prioritätenvergabe unverzüglich mit sofortiger Wirkung durchgeführt. Falls es nicht erteilt wird, bleibt die die bisher festgeschriebene Abschnittsweise Zug- oder Anschlußpriorität bis zur nächsten planmäßigen Prioritätenvergabe bestehen. Für alle zukünftig planmäßig angeforderten Prioritäten werden die entsprechenden Störfallprioritäten ohne manuelle Freigaben vergeben.

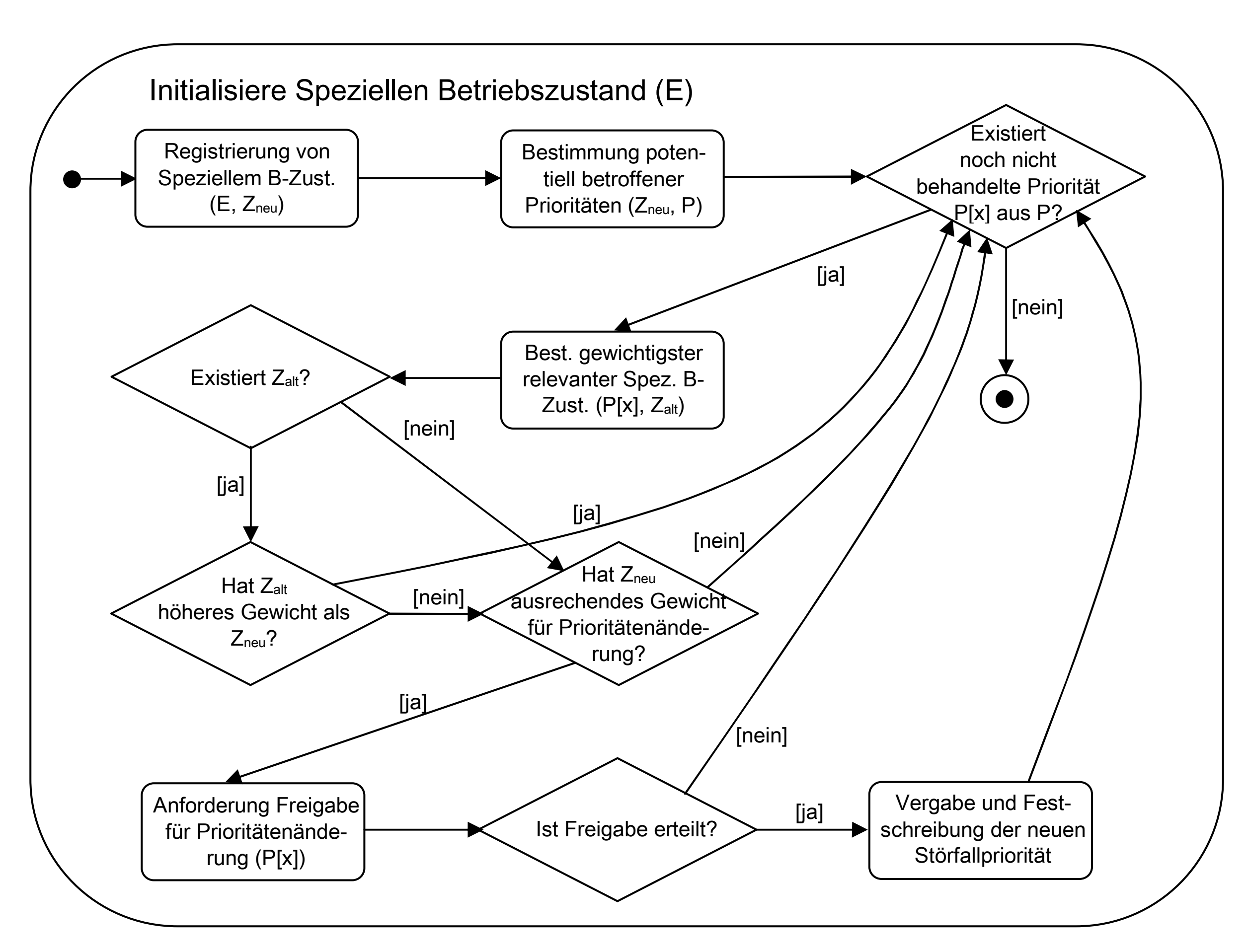


Abb. 6-10: Dynamisches Modell der Initialisierung eines Speziellen Betriebszustands

Das Verfahren für Ereignisse, die einen Speziellen Betriebszustand terminieren, ist sinngemäß dasselbe und soll daher nicht detailliert beschrieben werden (vgl. Abb. 6-11). Auf Darstellung von Modellen für Aktualisierungen von Speziellen Betriebszuständen wurde hier verzichtet. Sie lassen sich allerdings einfach aus den Initialisierungsmodellen herleiten. Für die Implementierung der störfallbedingten Prioritätenvergabe sowie für die Administration Spezieller Betriebszustände können aktive Regeln eingesetzt werden (vgl. 6.4.4 und 6.4.5).

Die **partielle Initialisierung von Standard- und Störfallprioritäten** entspricht weitestgehend der bereits in 6.3.1 beschriebenen vollständigen Prioritäteninitialisierung. Daher kann die Darstellung relativ kurz erfolgen und sich auf die Besonderheiten konzentrieren.

Im Falle von während des Systembetriebs nachträglich einzufügenden Zügen, geänderten Bonus-Malus-Regelungen und aktualisierten normativen Randbedingungen müssen einzelne Prioritäten (neu) initialisiert werden. Dabei soll sich das Prioritätenvergabe- und -optimierungssystem auf möglichst wenige, unbedingt neu zu initialisierende Prioritäten beschränken, um Optimierungsschritte der Vergangenheit nicht zu verlieren. Daher sollte ein komplettes Zurücksetzen nur in sehr seltenen Ausnahmefällen praktiziert werden, z. B. bei einer vollständigen Fahrplanrevision oder einem komplett neuen Bonus-Malus-System.

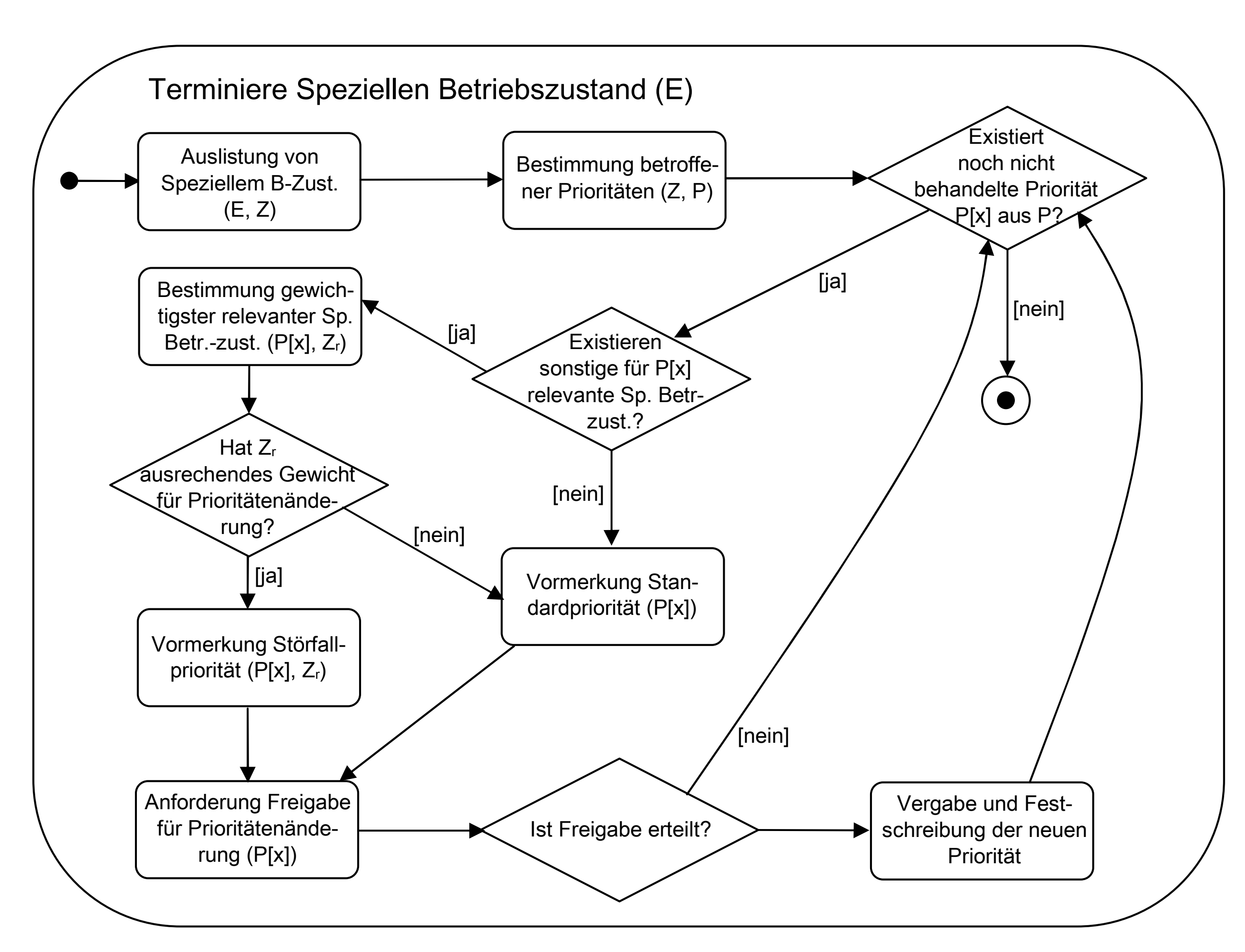


Abb. 6-11: Dynamisches Modell der Terminierung eines Speziellen Betriebszustandes

Wenn ein neuer Zug in das System eingefügt wird, müssen für alle Prioritätenabschnitte, auf denen der Zug verkehrt und für alle Anschlüsse, an denen der Zug beteiligt ist, Initialprioritäten bzgl. des neuen Zuges festgelegt werden. Dabei wird das unter 6.3.1 beschriebene Verfahren sinngemäß angewendet, ohne daß die übrigen Prioritäten geändert werden.

Wenn (kleinere) Änderungen an Bonus-Malus-Regelungen für eine Planabweichung oder einen Anschlußverlust vorgenommen werden, sollten für davon betroffene Prioritäten ebenfalls neu initialisiert werden. Bei der neubewerteten Planabweichung eines Zuges bezieht sich dies auf alle Abschnittsweisen Zug- und Anschlußprioritäten des Zuges vor Erreichen

des betroffenen Verkehrshaltes. Bei einer neuen Anschlußbewertung betrifft dies neben der Anschlußpriorität des Anschlusses selber alle Abschnittsweisen Zug- und Anschlußprioritäten des Zubringerzuges bis zum Erreichen des betroffenen Anschlußbahnhofs.

Bei einer Änderung der normativen Randbedingungen müssen alle neu eingeschränkten Prioritäten bzgl. ihrer „Legalität" überprüft werden. Wenn irgendwelche Abschnittsweisen Zug- oder Anschlußprioritäten einen nicht mehr zulässigen Wert aufweisen, muß dieser entsprechend angepaßt werden. Dabei muß ausnahmsweise nicht unbedingt das in 6.3.1 beschriebene Verfahren komplett angewendet werden. Es reicht aus, auf den „nächstgelegenen" legalen Wert zu korrigieren.

Im Prozeß der **Prioritätenoptimierung** steht die längerfristige Verbesserung der für Vergaben intern gespeicherten Standard- und Störfallprioritäten auf Basis des Bonus-Malus-Systems im Mittelpunkt. Das dahinterstehende Grundprinzip wurde bereits ausführlich in Kapitel 5 erläutert, so daß hier primär technische Aspekte der Abwicklung im Vordergrund stehen sollen.

Der Prozeß (vgl. Abb. 6-12) wird durch die Meldung von Bonus-Malus-Zahlungen seitens des Qualitätsüberwachungssystems ausgelöst. In Rahmen dieser Ereignismeldungen wird insbesondere mitgeteilt, in welchem Prioritätenbereich (nach- bzw. vorrangiger Zug in Prioritätenabschnitt, eingehaltener bzw. gebrochener Anschluß) die letztendliche Ursache für das ungünstige Ergebnis zu suchen ist, d. h. in welchem Netzbereich offensichtlich Handlungsbedarf besteht (vgl. Ausführungen in 5.2). Ein „hektisches" Reagieren auf jede noch so kleine Bonus-Malus-Zahlung soll im Sinne einer späteren Akzeptanz des Prioritätenvergabe- und -optimierungssystems in jedem Falle vermieden werden. Mit Hilfe der Definition von Toleranzgrenzen ist es möglich, mit etwas mehr „Gelassenheit" die Prioritätenoptimierung zu betreiben und nur auf wirklich nachhaltige, wiederholt auftretende Schwierigkeiten zu reagieren.

Die Bestimmung der Toleranzgrenzenüberschreitungen kann mittels deduktiver Regeln aufgrund der aktuellen und der vorausgegangenen Bonus-Malus-Zahlungen erfolgen (vgl. 6.4.3). Dabei werden allerdings ausschließlich jene Bonus-Malus-Ereignisse berücksichtigt, die seit der letzten Änderung der aktuell gültigen und an das Dispositionssystem vergebenen Standard- bzw. Störfallpriorität bzw. nach Optimierung von Prioritäten von „benachbarten" potentiellen Konfliktzügen aufgetreten sind. Damit erhalten kürzlich optimierte Prioritäten zunächst einmal eine gewisse Bewährungszeit, bevor weitere Optimierungsschritte begonnen werden können. Ferner wird als Untersuchungszeitraum ausschließlich die Anzahl der Betriebstage betrachtet, bei der die aktuell vergebene und unveränderte Standard- oder Störfallpriorität auch tatsächlich ebenfalls an das Dispositionssystem vergeben wurde, was insbesondere bei selten auftretenden Speziellen Betriebszuständen von Bedeutung ist.

Beim Eintreten einer Toleranzgrenzenüberschreitung kann die eigentliche Optimierung mit Hilfe von aktiven Regeln erfolgen, die durch entsprechende Überschreitung der o. g. Toleranzgrenzen für Verursacher von Bonus-Malus-Zahlungen aktiviert werden sollen und daraufhin entsprechende Anpassungen der gespeicherten Standard- und Störfallprioritäten vorschlagen (vgl. 6.4.6). Dabei gilt das Prinzip, daß beim gleichzeitigen Überschreiten mehrerer

Toleranzgrenzen die „gravierendste“ zunächst behandelt werden soll. Es erfolgt der Vorschlag zur Absenkung oder Anhebung einer Priorität entsprechend den Ausführungen in 5.2 und 5.3 inkl. des Resets von betroffenen Toleranzgrenzenüberwachungen.

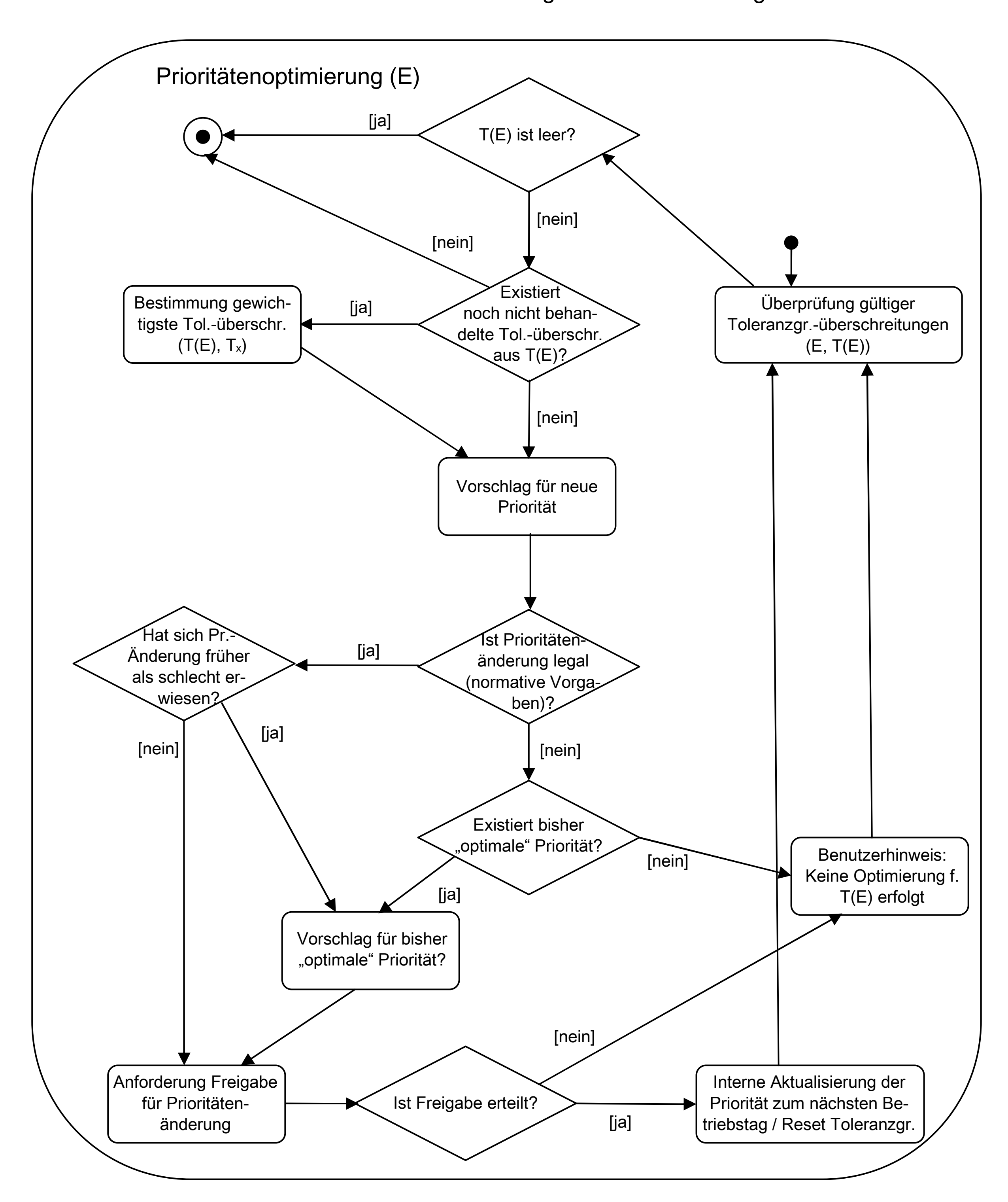


Abb. 6-12: Dynamisches Modell der Prioritätenoptimierung

Danach muß zunächst überprüft werden, ob die angestrebte Änderung mit den vorhandenen normativen Vorgaben vereinbar ist (Wertebereich der Prioritäten noch „legal“), was mit Hilfe normativer Regeln realisiert werden kann (vgl. 6.4.1). Falls dies der Fall ist, wird kontrolliert, ob die resultierende lokale Prioritätenkonstellation bereits in der Vergangenheit verwendet

wurde und sich in den Ergebnissen als ungünstig herausgestellt hat. Bei „Illegalität“ der Änderung oder Ungünstigkeit wird das bisherige lokale Optimum als Alternativlösung vorgeschlagen.

Vor der endgültigen Freischaltung einer Priorität kann eine Freigabe des Benutzers angefordert werden. Dauerhafte Prioritätenänderungen sollen nicht im Hintergrund ohne Wissen des Benutzers erfolgen. Bei Verweigerung wird die Optimierung zunächst abgebrochen, und die bestehenden Prioritäten bleiben vorerst in Kraft. Ansonsten wird die Änderung mit Wirkung vom nächsten Betriebstag an durchgeführt. Eine unverzügliche Freischaltung soll vermieden werden, um kurzfristige Änderungen im laufenden Betrieb zu vermeiden.

6.4 Regelbasierte Vergabe und Optimierung von Prioritäten

Bei der Entwicklung des Prioritätenvergabe- und -optimierungssystems sollen wesentliche Prozesse durch Regeln realisiert werden. Normative Regeln werden dazu verwendet, legale bzw. illegale Prioritätenvergaben aufgrund politischer oder betrieblich-organisatorischer Vorgaben definieren zu können. Deduktive Regeln unterstützen den Prioritätenvergabeprozeß für planmäßige Prioritäten durch Ableitung der jeweils geeigneten Standard- oder Störfallpriorität und durch Ableitung von den Optimierungsprozeß auslösenden Toleranzgrenzen für Bonus-Malus-Zahlungen bzw. zur Bewertung von Prioritätenkonstellationen. Aktive Regeln schließlich automatisieren Einfügung, Aktualisierung und Löschung Spezieller Betriebszustände durch Betriebsereignisse, die damit verbundene planmäßige und störfallbedingte Prioritätenvergabe und für die eigentliche Prioritätenoptimierung.

Im folgenden sollen nun die wesentlichen Einsatzgebiete der Regeln genauer betrachtet werden. Hierbei werden zur besseren Lesbarkeit der Regeln die im Abkürzungsverzeichnis im Anhang zu dieser Arbeit aufgeführten Kürzel als Variablen verwendet.

6.4.1 Kontrolle der Prioritätenvergabe durch normative Regeln

Ziel des Einsatzes normativer Regeln im Rahmen des Prioritätenvergabe- und -optimierungssystem ist die Einschränkung des Wertebereichs möglicher Standard- und Störfallprioritäten aufgrund politischer Rahmenbedingungen bzw. sonstiger grundsätzlicher Vorgaben des Infrastrukturbetreibers oder einer Regulierungsbehörde. Bei der Vergabe und Optimierung der Prioritäten durch das System müssen die Vorgaben der normativen Regeln grundsätzlich bei der Auswahl der Lösung beachtet werden. Es wird separat nach Abschnittsweisen Zug- und Anschlußprioritäten unterschieden.

Eine normative Regel für Abschnittsweise Zugprioritäten soll beispielhaft entwickelt werden. Das Schema für DB-Fakten, die Abschnittsweise Zugprioritäten repräsentieren, hat entsprechend des in Kapitel 4 entwickelten Objektmodells die folgende Form:

```
abs_zugpr(Bt,Mz,Pabs,Rg).
```

Dieses sowie später folgende Schemata lassen sich aus den entsprechenden ER-Diagrammen ableiten. Das dargestellte Fakt kann einem Entity mit seinen in Klammern angegebenen Attributen entsprechen oder es repräsentiert eine Verknüpfung mit den entsprechenden Schlüsselattributen der verknüpften Entities bzw. den Verknüpfungsattributen. Aus Gründen der Kompaktheit werden bei den Regeldarstellungen z. T. Abkürzungen bei den Bezeichnungen der Entity- bzw. Verknüpfungstypen vorgenommen.

Üblicherweise werden die Prioritäten für bestimmte Züge auf einem bestimmten Prioritätenabschnitt auf einen bestimmten Wertebereich eingeschränkt. Praktisch läßt sich dies einfach durch Ausschluß der nicht zulässigen Wertbereiche realisieren. Wenn für einen Modellzug mz1 im Prioritätenabschnitt pabs1 an allen Betriebstagen ausschließlich eine Rangstufe zwischen 2 und 5 zulässig sein soll, kann dies folgendermaßen modelliert werden:

```
unzul_rg_mz1_pabs1 ← abs_zugpr(_,mz1,pabs1,1).
unzul_rg_mz1_pabs1 ← abs_zugpr(_,mz1,pabs1,Rg), Rg > 5.
constraint not unzul_rg_mz1_pabs1.
```

Statt für einzelne Modellzüge bzw. einzelne Prioritätenabschnitte können die normativen Regeln auch für Mengen von Zügen oder Prioritätenabschnitte aufgestellt werden. So kann beispielsweise eine Vorrangstrecke für den Güterverkehr für einen Prioritätenabschnitt pabs2 definiert werden, bei denen die Standardprioritäten für Güterzüge zwischen 2 und 5 und die Standardprioritäten für Reisezüge größer als 6 sind. (Niedrige numerische Werte entsprechen einem hohen Rang des Zuges auf dem Prioritätenabschnitt):

```
unzul_nachrg_gz_pabs2 ← abs_zugpr_stand(_,Mz,pabs2,1), Mz.RzGz = Gz.
unzul_nachrg_gz_pabs2 ← abs_zugpr_stand(_,Mz,pabs2,Rg), Mz.RzGz = Gz, Rg > 5.
constraint not unzul_nachrg_gz_pabs2.
```

Für die Störfallprioritäten sollen in diesem Fall derartige Einschränkungen nicht gelten. In diesem Fall kann das gesamte Spektrum für alle Züge genutzt werden, um geeignet mit den Besonderheiten der Speziellen Betriebszustände umgehen zu können.

Bei Anschlußprioritäten können auf ähnliche Weise unzulässige Werte festgelegt werden. Fakten für Anschlußprioritäten haben die Form

```
anschl_pr(Bt,Anschl,Max_Wz).
```

So kann z. B. ein Qualitätsanschluß für den Anschluß anschl1 erzwungen werden, indem die maximale Wartezeit zwischen 5 und 15 Minuten festgelegt wird:

```
unzul_anschl_pr_anschl1 ← anschl_pr(_,anschl1,Wz), Wz < 5.
unzul_anschl_pr_anschl1 ← anschl_pr(_,anschl1,Wz), Wz > 15.
constraint not unzul_anschl_pr_anschl1.
```

Bei der Vergabe „harter“ normativer Regeln muß sehr vorsichtig agiert werden. Die normativen Regeln haben keine Rangstufen und müssen daher auch grundsätzlich (unabhängig von weiter definierten normativen Regeln) zu jedem Zeitpunkt beachtet werden. Wenn z. B. für langsame Züge nur eine Prioritätenmenge A zulässig ist, für Güterzüge jedoch nur eine Menge B, reduziert sich die Menge der zulässigen Prioritäten für „langsame Güterzüge“ automatisch auf die unter Umständen leere Schnittmenge von A und B. Um solche Effekte zu

vermeiden, sollte von normativen Regeln nur zurückhaltend Gebrauch gemacht werden. Man sollte sich auf echte politische Vorgaben konzentrieren, wo die Verwendung normativer Regeln unvermeidbar ist.

Alternativ besteht auch die Möglichkeit, „weiche“ normative Regeln festzulegen, die in Einzelfällen durch Freigabe des Benutzers gebrochen werden dürfen. Hier stehen insbesondere betrieblich wünschenswerte Vorgaben im Vordergrund, die aber nicht zwangsläufig in jedem Falle gehalten werden müssen.

6.4.2 Deduktive Regeln zur Ableitung von Prioritäten

Planmäßige Prioritätenvergaben werden durch das Dispositionssystem angefragt bzw. verbindlich angefordert, sobald sie für die weitere Disposition von Bedeutung sind. Dies ist für Abschnittsweise Zugprioritäten spätestens dann der Fall, wenn der betreffende Zug bereit für die Einfahrt in den neuen Prioritätenabschnitt ist. Für Anschlußprioritäten ist dies der Fall, sobald der potentiell anschlußaufnehmende Zug den Anschlußbahnhof erreicht.

Dem Prioritätenvergabe- und -optimierungssystem liegen sämtliche derzeit gültigen Standard- und Störfallprioritäten für Abschnittsweise Zug- und Anschlußprioritäten vor. Ferner sind die zum aktuellen Zeitpunkt aktiven Speziellen Betriebszustände im System gespeichert. Mittels deduktiver Regeln kann aus diesen Informationen einfach die jeweils gültige und zu vergebene Abschnittsweise Zug- und Anschlußpriorität ermittelt werden.

Prinzipiell kann folgendermaßen vorgegangen werden: Zunächst einmal muß überprüft werden, ob ein für die entsprechend angeforderte Priorität relevanter Spezieller Betriebszustand vorliegt. Bei mehreren parallel gültigen Speziellen Betriebszuständen ist ausschließlich der Zustand mit dem höchsten Gewicht des Typs maßgebend. Falls keine Störfallpriorität gilt, ist die Standardpriorität zu vergeben.

Grundsätzlich sollte man beim Entwurf des Systems bemüht sein, die Anzahl der für eine Abschnittsweise Zug- oder Anschlußpriorität definierten Speziellen Betriebszustandstypen bzw. Störfallprioritäten überschaubar zu halten. Störfallprioritäten sollen nur für diejenigen Fälle definiert werden, in denen betriebliche Umstände eine wirkliche Abweichung von den standardmäßig vergebenen Prioritäten erfordern, um Bonus-Malus-Zahlungen zu vermeiden bzw. reduzieren. Dazu gehören entsprechend primär bzgl. der Gefährdungen zur Auslösung von Bonus-Malus-Zahlungen und eingeschränkte Ressourcen. Dies erhöht die Wartbarkeit des Systems und vermeidet ein unerwünschtes häufiges Wechseln der Prioritäten.

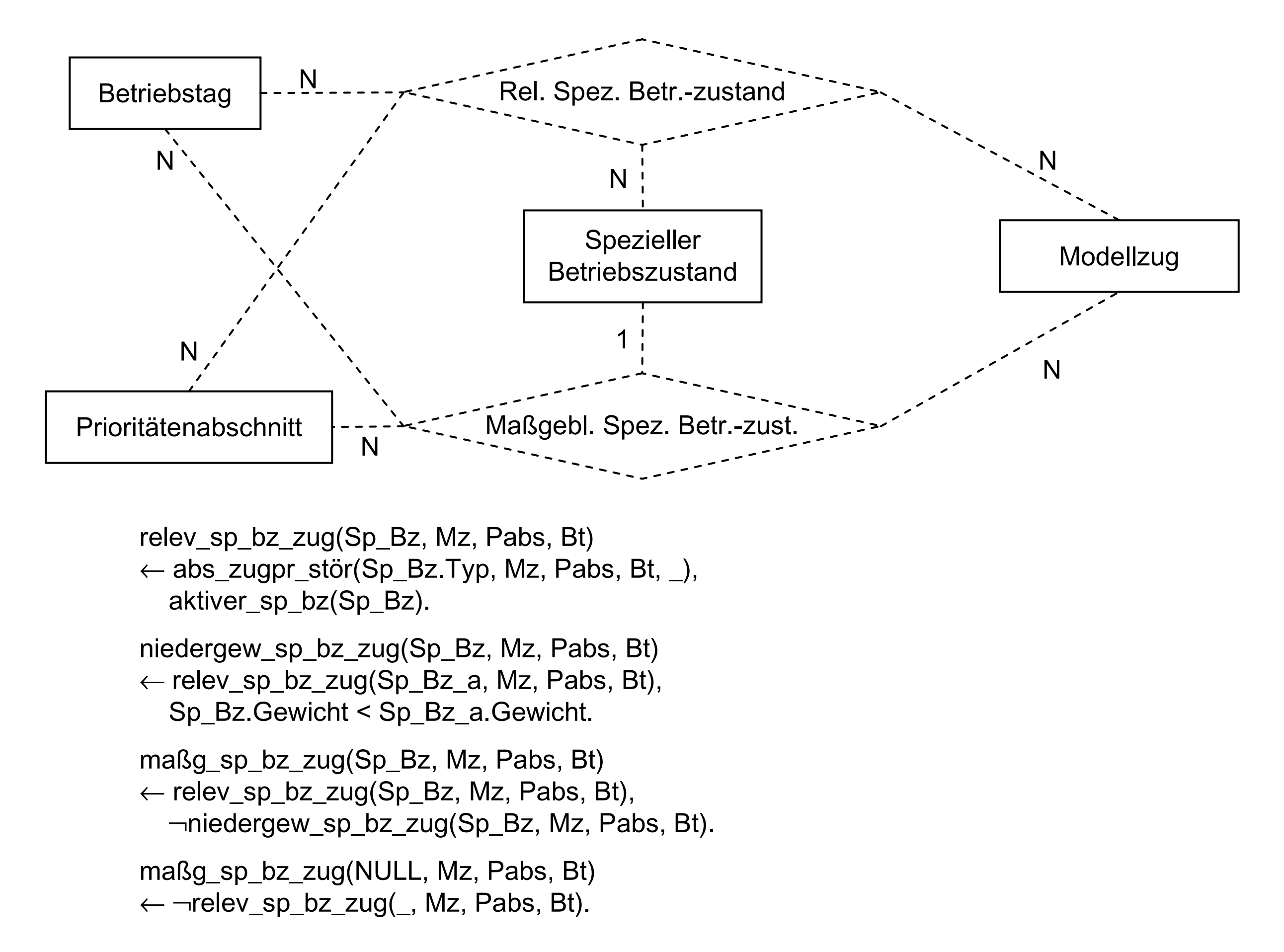
Im folgenden wird eine beispielhafte Formulierung für Abschnittsweise Zugprioritäten angegeben. Dazu werden folgende Relationen verwendet:

```
abs_zugpr_stör(Sp_Bz_Typ, Mz, Pabs, Bt, Rg).
abs_zugpr_stand(Mz, Pabs, Bt, Rg).
aktiver_sp_bz(Sp_Bz).
```

Da eine Störfallpriorität gleichzeitig mit mehreren Speziellen Betriebszustandstypen assoziiert sein kann, müßte eigentlich noch ein Identifikator der Störfallpriorität in die entsprechen-

de Relation aufgenommen werden. Aus Vereinfachungsgründen soll hier darauf verzichtet werden. Im Rahmen der Prioritätenoptimierung ist es aber von Bedeutung, daß alle verknüpften Speziellen Betriebszustandstypen die gleiche Rangzuordnung erhalten. Die letzte Relation ist eine Vereinigung aller im System aktuell (zum Zeitpunkt der Prioritätenvergabe) gültigen Speziellen Betriebszustände. Abgeleitet werden relevante und maßgebliche Spezielle Betriebszustände sowie die letztendlich gültige Abschnittsweise Zugpriorität (vgl. Abb. 6-16 und 6-17). Falls kein Maßgeblicher Spezieller Betriebszustand vorliegt, erhält das entsprechende Attribut den Wert NULL.

Für Anschlußprioritäten sind entsprechend analoge Formulierungen möglich.




```
relev_sp_bz_zug(Sp_Bz, Mz, Pabs, Bt)
← abs_zugpr_stör(Sp_Bz.Typ, Mz, Pabs, Bt, _),
   aktiver_sp_bz(Sp_Bz).

niedergew_sp_bz_zug(Sp_Bz, Mz, Pabs, Bt)
← relev_sp_bz_zug(Sp_Bz_a, Mz, Pabs, Bt),
   Sp_Bz.Gewicht < Sp_Bz_a.Gewicht.

maßg_sp_bz_zug(Sp_Bz, Mz, Pabs, Bt)
← relev_sp_bz_zug(Sp_Bz, Mz, Pabs, Bt),
   ¬niedergew_sp_bz_zug(Sp_Bz, Mz, Pabs, Bt).

maßg_sp_bz_zug(NULL, Mz, Pabs, Bt)
← ¬relev_sp_bz_zug(_, Mz, Pabs, Bt).
```

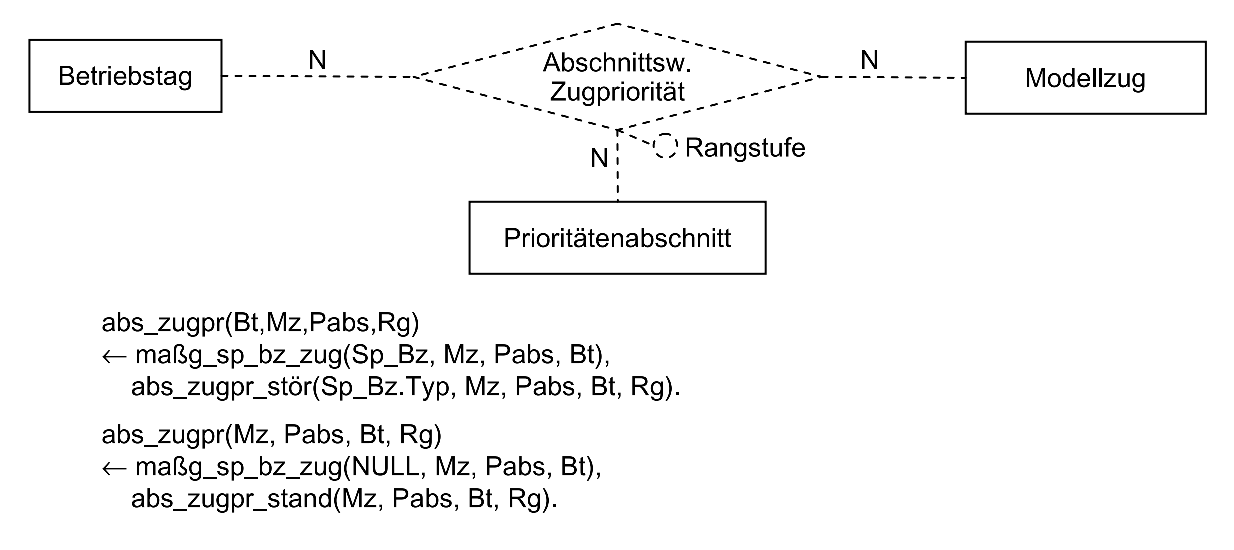
Abb. 6-16: Relevante und maßgebliche Spezielle Betriebszustände

Als nachteilig bei dieser Modellierung stellt sich allerdings heraus, daß relevante und maßgebliche Betriebszustände wegen der Ableitung aus aktiver_sp_bz(…) ausschließlich im aktuellen Zeitraum angegeben werden können. Eine Rückschau auf maßgebliche Betriebszustände zu einem Zeitpunkt in der Vergangenheit ist so nicht möglich. Da es alleine schon wegen des Datenvolumens nicht zielführend ist, sämtliche einmal zu irgendeinem Zeitpunkt im System gültigen Speziellen Betriebszustände zu archivieren, macht es Sinn, zumindest die zum Zeitpunkt der Prioritätenvergaben für diese Priorität jeweils maßgeblichen Speziellen Systemzustände durch das Prioritätenvergabe- und -optimierungssystem mittelfristig in einer Relation

maßg_sp_bz_zug_verg(…). bzw.
maßg_sp_bz_anschl_verg(…).

zu speichern.

Damit kann die an jedem Betriebstag verwendete Standard- bzw. Störfallpriorität zu einem späteren Zeitpunkt nachvollzogen werden.



abs_zugpr(Bt,Mz,Pabs,Rg)
← maßg_sp_bz_zug(Sp_Bz, Mz, Pabs, Bt),
abs_zugpr_stör(Sp_Bz.Typ, Mz, Pabs, Bt, Rg).

abs_zugpr(Mz, Pabs, Bt, Rg)
← maßg_sp_bz_zug(NULL, Mz, Pabs, Bt),
abs_zugpr_stand(Mz, Pabs, Bt, Rg).

Abb. 6-17: Abgeleitete aktuelle Abschnittsweise Zugpriorität

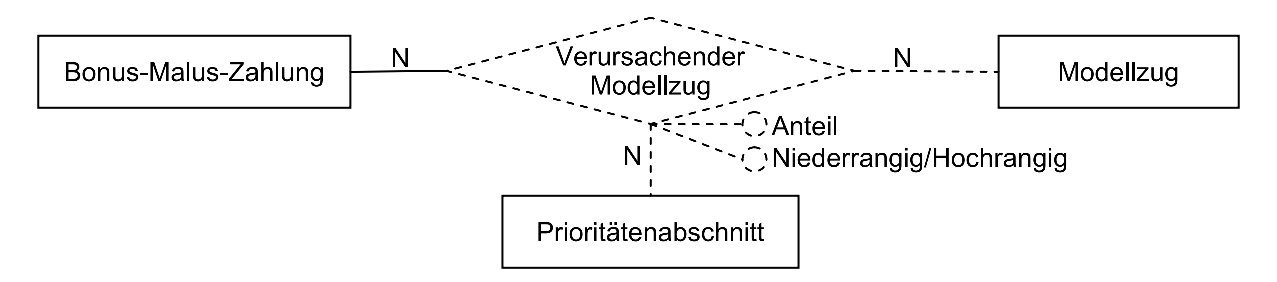
6.4.3 Deduktive Regeln zur Bestimmung von Toleranzgrenzenüberschreitungen

Im Rahmen des Eisenbahnbetriebs fallen regelmäßig Planabweichungen in Form von Abweichungen bei Ankunfts- bzw. Abfahrtszeiten oder gebrochenen Anschlüssen an. Sie werden vom Qualitätsüberwachungssystem in Form von Bonus-Malus-Zahlungen signalisiert. Diese Planabweichungen werden entweder in Prioritätenabschnitten oder Anschlüssen verursacht. Für die verschiedenen Verursacher gibt es pro Modellzug oder Anschluß verschiedene Toleranzgrenzen zulässiger Bonus-Malus-Zahlungen in einem vorgegebenen Zeitraum, jeweils spezifiziert für relativ hochrangige und niederrangige Züge bzw. eingehaltene und gebrochene Anschlüsse. Sobald mehr Bonus-Malus-Zahlungen als tolerabel durch einen Verursacher ausgelöst werden, muß diese Toleranzgrenzenüberschreitung zunächst einmal vom Prioritätenvergabe- und -optimierungssystem geeignet erkannt werden. Dies kann mit Hilfe deduktiver Regeln erfolgen.

Eine Toleranzgrenzenüberschreitung ist immer identisch mit einem Bonus-Malus-Ereignis. Zusätzlich müssen vergangene Zahlungen betrachtet werden, wobei nur Zahlungen seit dem letzten Reset-Zeitpunkt für die Toleranzgrenze berücksichtigt werden dürfen. Dazu müssen die relevanten Betriebstage bestimmt werden, die für die Toleranzgrenzenüberschreitung betrachtet werden müssen. Dabei handelt es sich um jene in der Toleranzgrenze festgelegten nach dem letzten Reset-Zeitpunkt liegenden letzten n Betriebstage, an denen der Modellzug den Prioritätenabschnitt befahren bzw. der Anschluß geplant wurde, bei dem die aktuell verwendete Standard- bzw. Störfallpriorität ebenfalls verwendet und keine Ände-

rung von deren Wert vorgenommen wurde. Dies soll nun beispielhaft für die in einem Prioritätenabschnitt relativ niederrangigen Züge regelbasiert umgesetzt werden.

In 5.2.1 wurde gezeigt, wie Bonus-Malus-Zahlungen auf einzelne prioritätenbasierte Dispositionsmaßnahmen anteilmäßig aufgeteilt werden können. Daraus läßt sich die folgende Relation intuitiv ableiten (hier ohne Angabe der Ableitungsregeln):

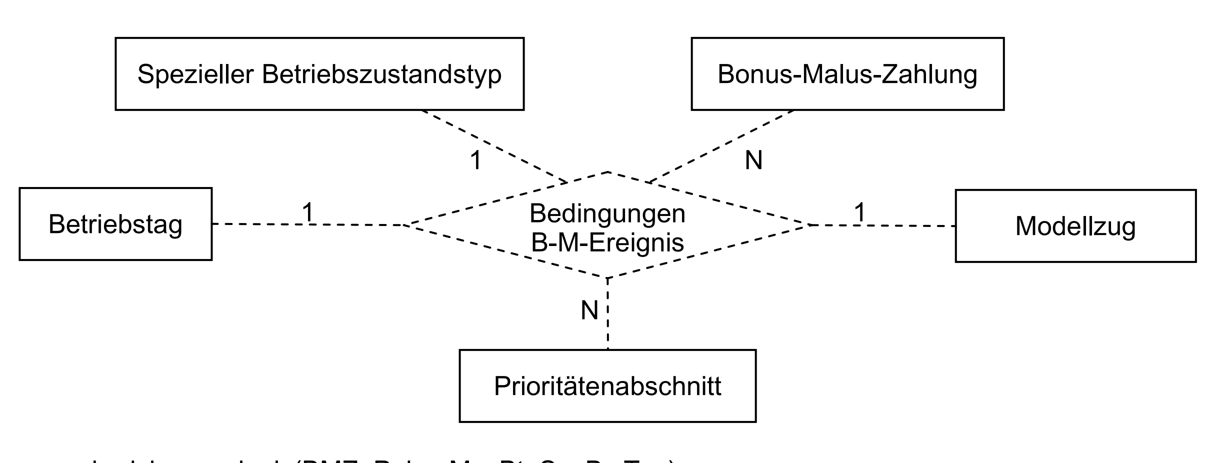


verurs_mz(BMZ, Pabs, NrHr, Mz, Anteil).

Abb. 6-18: Verursachende Modellzüge

Weiterhin sollte bestimmt werden, ob auf dem die Bonus-Malus-Zahlung verursachenden Prioritätenabschnitt Standard- oder Störfallbedingungen vorhanden waren, indem der maßgebliche Spezielle Betriebszustand bei Vergabe (so vorhanden) bestimmt wird. Ferner sollen alle wesentlichen Attribute der Bonus-Malus-Zahlung in einer Relation gebündelt werden (siehe Abb. 6-19). Die zugehörigen Relationen sind:

bm_bzgl(BMZ, Pabw).
pabw_datum(Pabw, Bt).

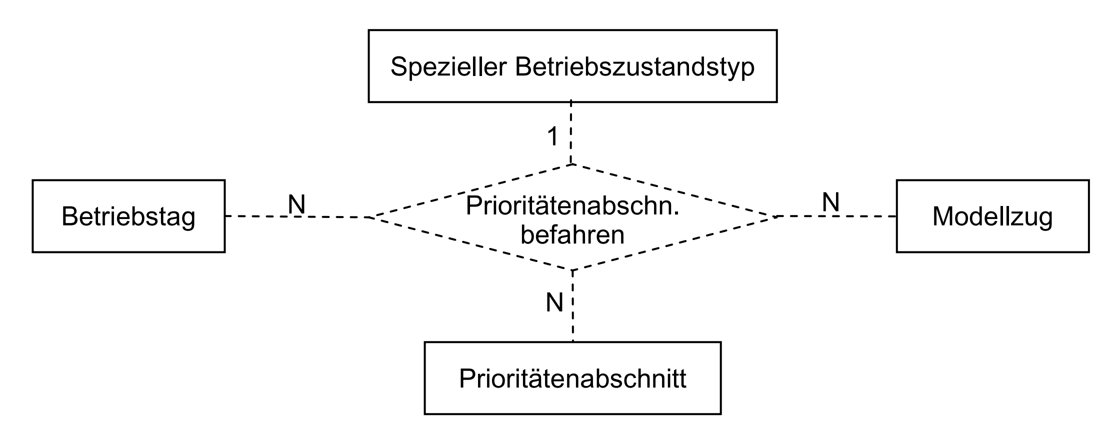


bed_bm_ereignis(BMZ, Pabs, Mz, Bt, Sp_Bz.Typ)
← verurs_mz(BMZ, Pabs, _, Mz, _),
bm_bzgl(BMZ, Pabw),
pabw_datum(Pabw, Bt),
maßg_sp_bz_zug_verg (Sp_Bz, Mz, Pabs, Bt).

Abb. 6-19: Bedingungen bei Bonus-Malus-Ereignis

Ferner müssen die Betriebstage bestimmt werden, an denen der Modellzug einen bestimmten Prioritätenabschnitt befahren hatte. Dazu sollte der Typ des dazu maßgeblichen Speziellen Betriebszustands als Attribut gelistet werden (siehe Abb. 6-20). Dies läßt sich aus der

Streckenabschnittsliste des Modellzuges, der Prioritätenabschnittsdefinition und dem maßgeblichen Speziellen Betriebszustand ableiten. Auf die Darstellung der Regel wird verzichtet.

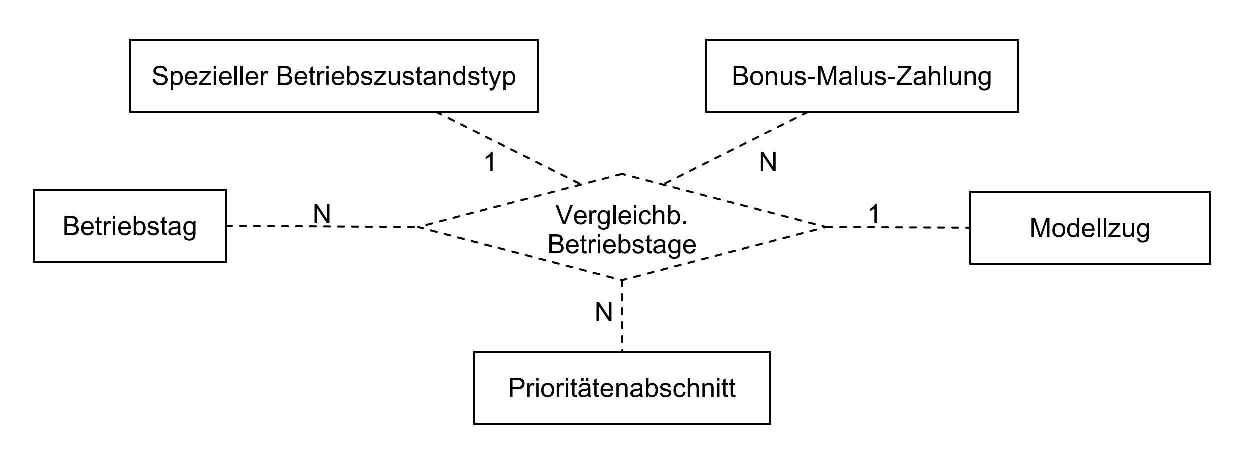


pabs_bef(Pabs, Mz, Bt, Sp_Bz_Typ).

Abb. 6-20: Befahrene Prioritätenabschnitte

Im folgenden soll sollen alle Betriebstage nach dem letzten Reset-Datum bestimmt werden, an denen der Zug auf dem Prioritätenabschnitt mit derselben Standard- oder Störfallpriorität wie bei der Bonus-Malus-Zahlung verkehrt ist (siehe Abb. 6-21). Dies basiert auf der selbsterklärenden Relation

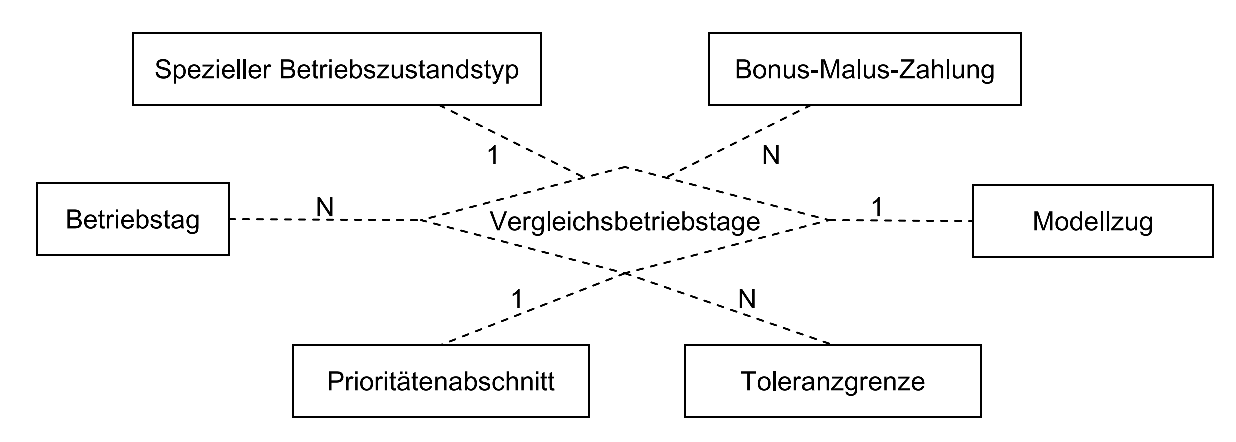
prior_reset(Pabs, Mz, Bt, Sp_Bz_Typ).



vglb_bt(BMZ, Pabs, Mz, Sp_Bz_Typ, Relev_Bt)
← bed_bm_ereignis(BMZ, Pabs, Mz, _, Sp_Bz_Typ),
pabs_bef(Pabs, Mz, Relev_Bt, Sp_Bz_Typ),
prior_reset(Pabs, Mz, Reset_Bt, Sp_Bz_Typ),
Relev_Bt > Reset_Bt,
¬prior_reset(Pabs, Mz, Relev_Bt, Sp_Bz_Typ).

Abb. 6-21: Vergleichbare Betriebstage

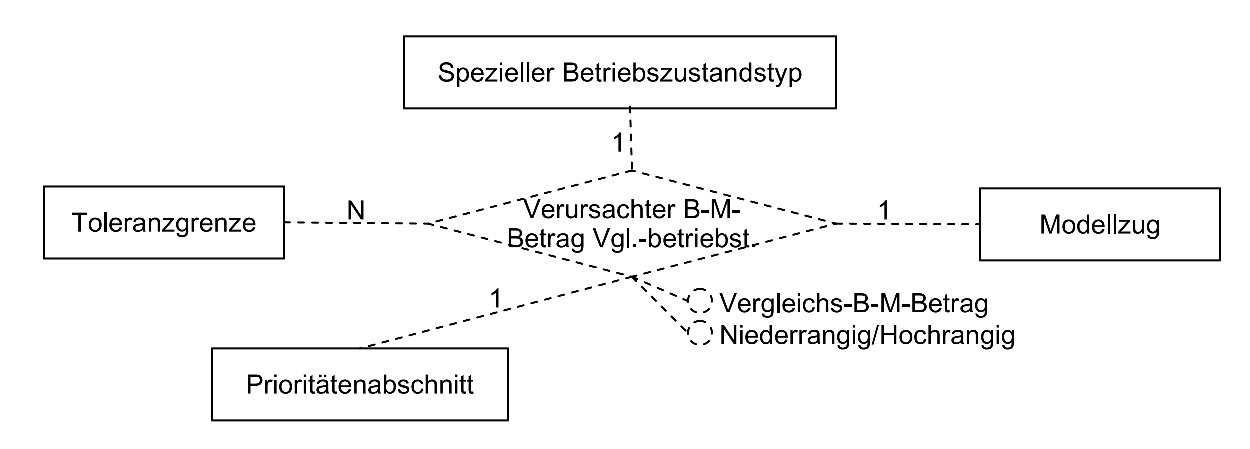
Hieraus können die für eine Toleranzgrenzenüberwachung relevanten Vergleichsbetriebstage ermittelt werden. Hierbei handelt es sich um die letzten n vergleichbaren Betriebstage, wobei es sich bei n um den zu überwachenden Toleranzgrenzenzeitraum handelt (siehe Abb. 6-22). Auf die Formulierung der Ableitungsregeln wird wiederum an dieser Stelle verzichtet.



vgl_bt(BMZ, Pabs, Mz, Sp_Bz_Typ, Tg, VglBt).

Abb. 6-22: Vergleichsbetriebstage

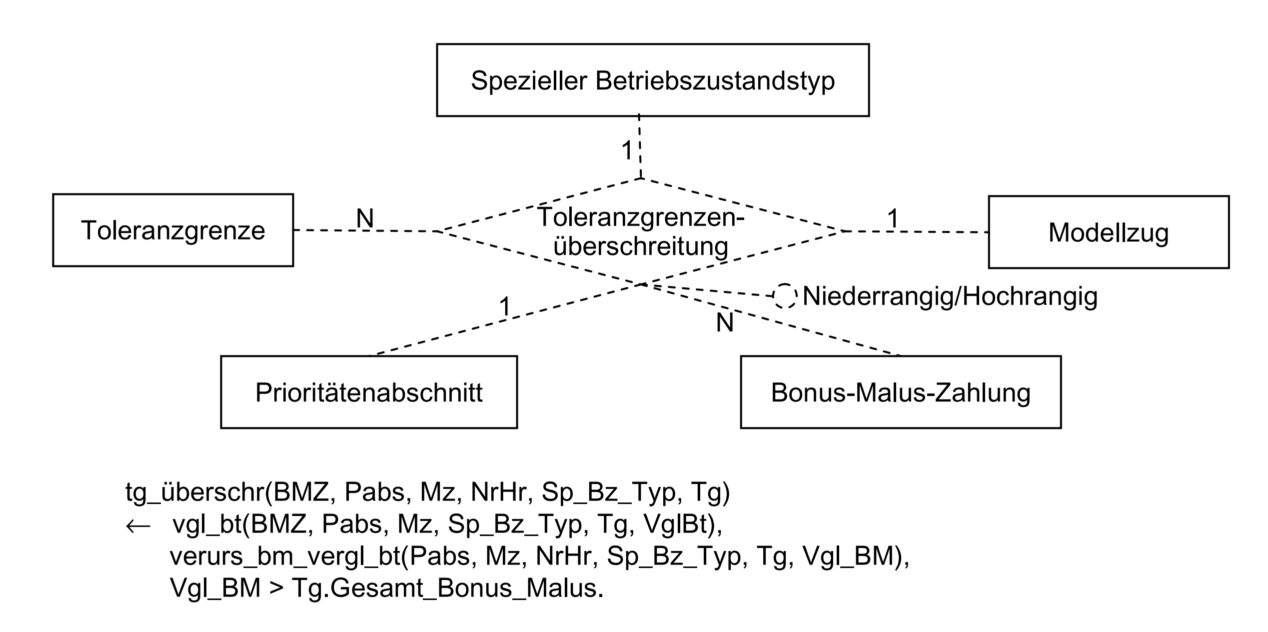
Für diese Vergleichsbetriebstage erfolgt nun die eigentliche Toleranzgrenzenprüfung. Hierbei muß die Summe aller für den Modellzug auf dem Prioritätenabschnitt verursachten Bonus-Malus-Zahlungen über den Betrachtungszeitraum einer vorgegeben Toleranzgrenze berechnet werden. Hierbei werden sämtliche innerhalb der Vergleichsbetriebstage anfallenden und von dem Modellzug zumindest teilweise entsprechend der Schuldanteile verursachte Bonus-Malus-Zahlungen aufgrund von Planabweichungen durch Verspätungen oder verpaßte Anschlüsse (mit dem Modellzug als verspäteten Anschlußzubringer) zusammengefaßt (siehe Abb. 6-23, hier wieder ohne Regeldarstellung).



verurs_bm_vergl_bt(Pabs, Mz, NrHr, Sp_Bz_Typ, Tg, Vgl_BM).

Abb. 6-23: Verursachter Bonus-Malus-Betrag im Vergleichszeitraum für Toleranzgrenze

Schließlich kann mit diesen Informationen einfach abgeleitet werden, ob durch eine Bonus-Malus-Zahlung eine Toleranzgrenzenverletzung bzgl. durch einen Zug in einem Prioritätenabschnitt verursachter Bonus-Malus-Anteile ausgelöst wird (vgl. Abb. 6-24).



tg_überschr(BMZ, Pabs, Mz, NrHr, Sp_Bz_Typ, Tg)
← vgl_bt(BMZ, Pabs, Mz, Sp_Bz_Typ, Tg, VglBt),
verurs_bm_vergl_bt(Pabs, Mz, NrHr, Sp_Bz_Typ, Tg, Vgl_BM),
Vgl_BM > Tg.Gesamt_Bonus_Malus.

Abb. 6-24: Toleranzgrenzenüberschreitung

Analog können ähnliche Prüfungen für durch Anschlußaufnahmen/Anschlußprioritäten verursachte Bonus-Malus-Zahlungen implementiert werden.

Grundsätzlich können durch eine Bonus-Malus-Zahlung gleichzeitig mehrere Toleranzgrenzen für einen oder mehrere Modellzüge auf einem Prioritätenabschnitt unter einem Speziellen Betriebszustand überschritten werden. Dementsprechend kann analog zu Speziellen Betriebszuständen eine entsprechend des verursachten Bonus-Malus-Anteils gewichtigste Toleranzgrenzenüberschreitung für eine Bonus-Malus-Zahlung auf einem Prioritätenabschnitt abgeleitet werden. Dies soll sicherstellen, daß für dasselbe Ereignis nicht mehrmals für denselben oder miteinander bzgl. Dispositionsmaßnahmen möglicherweise in Konflikt befindliche Züge reagiert wird.

Weitere, im Rahmen dieser Arbeit nicht mehr darstellbare Einsatzgebiete für deduktive Regeln sind im Rahmen des entwickelten Prioritätenvergabe- und -optimierungssystem möglich. So kann ein strukturell ähnliches Verfahren für die längerfristige Auswertung und Vergleiche von relativen lokalen Prioritätenkonstellationen gemäß 5.3 verwendet werden. Auf diese Weise läßt sich auch im Falle von Prioritätenanpassungen für einen Modellzug die Menge von Modellzug / Spezieller Betriebszustandstyp-Kombinationen bestimmen, die mit dem Modellzug in Vergangenheit in Konflikt bzgl. prioritätenbasierten Dispositionsmaßnahmen stand und daher von Reset-Ereignissen für Toleranzgrenzenüberschreitungen betroffen ist. Ferner lassen sich im Falle von Absenkungen und Anhebungen von Prioritäten gemäß den Ausführungen in 5.2 neue Prioritätenstufen ermitteln (die gleichzeitig von der normativen Komponente bzgl. Legalität überwacht werden).

6.4.4 Aktive Regeln zur Verwaltung Spezieller Betriebszustände

Bei den die Prioritätenvergabe beeinflussenden Speziellen Betriebszuständen handelt es sich in der Regel um bestimmte Gefährdungen von Bonus-Malus-Zahlungen verursacht durch Planabweichungen bzw. punktuelle Störungen oder um Ressourceneinschränkungen

durch kontinuierliche zug- oder fahrwegseitige Störungen. Ein Spezieller Betriebszustand wird durch ein vom Dispositionssystem gemeldetes initialisierendes Betriebsereignis (Planabweichungsmeldung eines Zuges für eine bestimmte Lokalität bzw. Terminierung einer punktuellen Störung oder Initialisierung einer kontinuierliche Störung) begonnen und ggf. durch ein korrespondierendes Änderungsereignis aktualisiert bzw. durch ein Terminierungsereignis beendet. Durch real stattgefundene Planabweichungen bzw. abgeschlossener punktuelle Störungen verursachte Spezielle Betriebszustände können verständlicherweise nicht terminiert werden, für lediglich prognostizierte Planabweichungen ist dies allerdings möglich. Allgemein kann man damit sagen, daß ein Spezieller Betriebszustand vorliegt, wenn ein initialisierendes Ereignis eingetroffen ist, aber das korrespondierende terminierende Ereignis noch nicht aufgetreten ist. Die Attribute des Speziellen Betriebszustandes werden durch das Initialisierungs- bzw. das zuletzt eingetretene Aktualisierungsereignis festgelegt.

Das Prioritätenvergabe- und -optimierungssystem muß die jeweils aktuellen Speziellen Betriebszustände intern verwalten, um im Einzelfall entscheiden zu können, ob bei Anforderung durch das Dispositionssystem die Standard- oder eine Störfallpriorität als Abschnittsweise Zug- oder Anschlußpriorität vergeben werden soll. Da jegliche Änderung im Zusammenhang mit Speziellen Betriebszuständen immer durch eine Betriebsereignismeldung des Dispositionssystems ausgelöst wird, eignen sich aktive Regeln in besonderer Weise zur Pflege der Speziellen Betriebszustände.

Die Grundmuster zur Einfügung, Aktualisierung und Terminierung von die Prioritätenvergabe beeinflussenden elementaren Speziellen Betriebszustände haben entsprechend die folgende generische Form:

on {Initialisierendes_Betriebsereignis(…)}

if {Erfüllung bestimmter für den Speziellen Betriebszustand relevanter Bedingungen und Spezieller Betriebszustand nicht im System vorhanden}

do {+ Spezieller Betriebszustand(…)}.

on {Aktualisierendes_Betriebsereignis(…)}

if { Erfüllung bestimmter für den Speziellen Betriebszustand relevanter Bedingungen und Relevanter Spezieller Betriebszustand im System vorhanden}

do {Verändere entsprechende Parameter des Speziellen Betriebszustandes}.

on {Terminierendes_Betriebsereignis(…)}

if { Nicht-Erfüllung bestimmter für den Speziellen Betriebszustand relevanter Bedingungen und Relevanter Spezieller Betriebszustand im System vorhanden}

do {– Spezieller Betriebszustand(…)}.

Die Zeichen + und – entsprechen im Datenbankkontext der Einfügung bzw. Löschung von Fakten. In diesem Falle geben sie die Initialisierung bzw. Terminierung für Spezielle Betriebszustände eines bestimmten Typs an, die entsprechend in der Datenbank gespeichert werden. Im folgenden sollen nun einzelne konkrete Beispielszenarien für mögliche Planabweichungen und Störungen vorgestellt werden.

Planabweichungen unterscheidet man in Abweichungen bzgl. Ankunfts- und Abfahrtszeiten sowie in Anschlußbrüche. Diese können für den Zug entsprechend der Bonus-Malus-Regelungen evtl. die Anwendung von Störfallprioritäten nahelegen. Bei geringen Verspätungen empfiehlt sich ggf. eine Höherbewertung des Zuges, um die Verspätung und damit evtl. das Risiko für Bonus-Malus-Zahlungen entsprechend zu reduzieren. Bei großen Verspätungen hingegen sollte der Zug ggf. herabgestuft werden, da sich die Bonus-Malus-Zahlung ohnehin selbst bei optimalem Betriebsverlauf nicht mehr verhindern läßt.

Es werden gemäß den Bonus-Malus-Regelungen für Verspätungen und Anschlußbrüche verschiede Intervallgrenzen für Verspätungen festgelegt, aus denen man dann Spezielle Betriebszustände zur Gefährdung von Bonus-Malus-Zahlungen ableiten kann. Ein Vorschlag für die Festlegung der Intervallgrenzen wird in 5.1 dargestellt. So könnten beispielsweise Spezielle Betriebszustandstypen für Verspätungen zwischen 15 und 35 Minuten (Gefährdung BM_1) sowie zwischen 30 und 60 Minuten (Gefährdung BM_2) definiert werden. Für einen Zug mit 32 Minuten Verspätung wären damit beide Spezielle Betriebszustände aktuell gültig. Dementsprechend müssen Spezielle Betriebszustandstypen mit einer höheren Steigerung der assoziierten Bonus-Malus-Zahlung mit einem höheren Gewicht als Zustände mit einer geringeren Steigerung versehen werden, damit diese gewichtigeren Zustände bei der Vergabe der Störfallprioritäten maßgeblich sind.

Störfallprioritäten für Abschnittsweise Zugprioritäten basierend auf Gefährdungen von Bonus-Malus-Zahlungen durch Planabweichungen sollen möglich sein im Falle von größeren Planabweichungen bei Ankunft im „Grenzbahnhof“ vor der Einfahrt in den jeweilig betroffenen Prioritätenabschnitt. Für Anschlußprioritäten sollen keine Störfallprioritäten basierend auf Planabweichungen definiert werden, da hier bereits die Zeitkomponente in der Prioritätendefinition selber enthalten ist.

Die Regel zur Einführung eines eine Störfallpriorität implizierenden Speziellen Betriebszustandes „Ankunftsverspätung (Vh, Bt, ProReal, Gew, 15, 35)“ für Ankunftsverspätungen zwischen 15 und 35 Minuten kann auf folgende Weise initialisiert werden. Das Gewicht des Speziellen Betriebszustandes soll in diesem Beispiel 5 sein.

```
on Planabw_VH(Zt, Vh, Bt, ankunft, Betrag, ProReal)

if Betrag >= 15,
   Betrag <= 35,
   not Ankunftsverspätung(Vh, Bt, _, 5, 15, 35)

do + Ankunftsverspätung(Vh, Bt, ProReal, 5, 15, 35).
```

Das einzige Attribut des Speziellen Betriebszustandes, das während seiner Gültigkeit aktualisiert werden kann, ist das Progn/Real-Attribut. Hier kann irgendwann eine Setzung des Wertes auf Real erfolgen. Daraus ergibt sich die folgende aktive Aktualisierungs-Regel für das obige Beispiel:

```
on Planabw_VH(Zt, Vh, Bt, ankunft, Betrag, real)

if Betrag >= 15,
   Betrag <= 35,
   Ankunftsverspätung(Vh, Bt, progn, 5, 15, 35)

do Ankunftsverspätung(Vh, Bt, _, 5, 15, 35).ProReal = real.
```

Eine Löschung dieses Speziellen Betriebszustandes ist nur für „prognostizierte“ Zustände möglich. Reale Planabweichungen können nicht mehr terminiert werden. Hier würden sich die folgende aktiven Regeln anbieten:

```
on Planabw_VH(Zt, Vh, Bt, ankunft, Betrag, ProReal)

if Betrag < 15,
   Ankunftsverspätung(Vh, Bt, progn, 5, 15, 35)

do – Ankunftsverspätung(Vh, Bt, progn, 5, 15, 35).

on Planabw_VH(Zt, Vh, Bt, ankunft, Betrag, ProReal)

if Betrag > 35,
   Ankunftsverspätung(Vh, Bt, progn, 5, 15, 35)

do – Ankunftsverspätung(Vh, Bt, progn, 5, 15, 35).
```

Derartige Regeln müssen für jeden durch Planabweichungen verursachten Speziellen Betriebszustandstyp eines Modellzuges in einem Prioritätenabschnitt (möglichst rechnerunterstützt) generiert werden.

Für abgeschlossene **punktuelle Störungen** kann ähnlich vorgegangen werden. Hier müssen allerdings nur Initialisierungsregeln für die Speziellen Betriebszustandstypen festgelegt werden, die aufgrund von Störungs-Terminierungen ausgelöst werden. Aktualisierungs- und Terminierungsregeln sind für diese Speziellen Betriebszustandstypen nicht erforderlich. Neben dem analog zu Planabweichungen überprüften Intervall der Gesamtverspätung soll außerdem die von der Störung ausgelöste Zusatzverspätung einen vorgegebenen Schwellenwert überschreiten, um auszuschließen, daß für kleine Störungen von wenigen Minuten bereits eine störfallbedingte Prioritätenvergabe ausgelöst wird. Man beachte, daß in der Regel für jeden vorgegebenen durch punktuelle Störungen verursachten Speziellen Betriebszustandstyp ein semantisch gleichwertiger durch eine Planabweichung verursachter Spezieller Betriebszustandstyp existiert, nämlich derjenige, der die Einhaltung derselben Bonus-Malus-Vorgabe gefährdet. Dabei soll das Gewicht des „Störungs-Speziellen Betriebszustand“ unterhalb dem des „Planabweichungs-Speziellen Betriebszustandes“ liegen, um die Eindeutigkeit des maßgeblichen Speziellen Betriebszustandes sicherzustellen. Auf genaue Syntaxbeispiele für Regeln im Zusammenhang mit durch punktuelle Störungen verursachte Speziellen Betriebszustände soll wegen der Ähnlichkeit zu Planabweichungen verzichtet werden.

Kontinuierliche Störungen sind ein Kerneinsatzgebiet für Störfallprioritäten. Hierbei stehen längerfristige Ressourceneinschränkungen im Vordergrund. Kurzfristige und abgeschlossene punktuelle Störungen werden bzgl. Störfallprioritäten über die oben genannten Speziellen Betriebszustandstypen abgedeckt. Man unterscheidet zwischen Zug- und Infrastrukturstörungen. Die Infrastrukturstörungen können je nach Auftrittsort der Störung weiter in Knoten- und Streckenstörungen unterteilt werden. Grundsätzlich sind im Eisenbahnbetrieb eine

Vielzahl an technischen, organisatorischen und sonstigen Störungen möglich. Um die Zahl der möglichen Speziellen Betriebszustandstypen und der damit verbundenen möglichen Störfallprioritäten in einem für die Anwendung im Zusammenhang mit Performanz und Wartbarkeit sinnvollem Umfang zu halten, sollte man hier auf eine begrenzte Zahl möglichst generischer Zustände zurückgreifen, die die Fälle einer ggf. benötigten störungsbedingten Prioritätenanpassung in ausreichender Genauigkeit abdecken. Genauer wird dies in 5.1 ausgeführt.

Entsprechend sollten die vom Dispositionssystem übermittelten störungsinitialisierenden, -aktualisierenden oder -terminerenden Betriebsereignisse diesem generischen Format genügen. Falls das aufgrund der Architektur des Dispositionssystems nicht möglich sein sollte, besteht die Möglichkeit, im Prioritätenvergabe- und -optimierungssystem auf einfache Weise eine Umwandlung in das generische Format mittels deduktiver Regeln durchzuführen. Zugstörungen können in diesem Falle beispielsweise generisch als eingeschränkte Höchstgeschwindigkeiten (hierunter fallen vor allem Motorenstörungen) oder verlängerte Abfertigungszeiten (hierunter fallen vor allem Störungen am Wagenmaterial oder bzgl. der Zugreihung) abgegrenzt werden.

Für streckenseitige Infrastrukturstörungen bieten sich einerseits Fahrtzeitverlängerungen (insbes. reduzierte Höchstgeschwindigkeiten auf Teilen der Strecke) und andererseits Kapazitätseinschränkungen (hierzu gehören vor allem Sperrungen einzelner Gleise auf Teilen der Strecke) als generische elementare Spezielle Betriebszustände an. Knotenseitige Infrastrukturstörungen sind für die Prioritätenvergabe in der Regel weniger bedeutsam. Als generische Störung eignen sich hier höchstens generelle Haltezeitverlängerungen (hierunter fallen z. B. längerfristige Signalstörungen oder im Güterverkehr technische Probleme in Rangierbahnhöfen).

Eine Besonderheit stellen extern verursachte Störungen dar. Eine Unterscheidung zwischen internen und externen Ursachen kann sinnvoll sein, wenn bei externen Ursachen („höhere Gewalt“) unter Umständen Bonus-Malus-Zahlungen entfallen und damit niedrigere Prioritäten angesetzt werden können. Daher kann in sollen Fällen ein Spezieller Betriebszustand vom Typ „Entfall aller Bonus-Malus-Zahlungen“ für einen Zug initialisiert werden. Dieser Zustand bleibt aktiviert, solange der betroffene Zug verkehrt.

Störfallprioritäten für Abschnittsweise Zugprioritäten können potentiell definiert werden für sämtliche Zugstörungen sowie für alle Infrastrukturstörungen im Bereich des jeweiligen Prioritätenabschnittes inkl. des „Anfangs-“ und „Endbahnhofes“ des Abschnittes. Störfallprioritäten für Anschlußprioritäten aufgrund von Zug- oder Streckenstörungen sind ebenfalls möglich. So kann auf Anschlußaufnahmen verzichtet werden, wenn der aufnehmende Zug im folgenden Prioritätenabschnitt eine deutlich verlängerte Fahrtzeit zu erwarten hat.

Die aktiven Regeln zur Initialisierung, Aktualisierung und Terminierung der durch kontinuierliche Störungen verursachten Speziellen Betriebszustände entsprechen syntaktisch weitestgehend denen der Planabweichungen und sollen daher hier nicht im Detail behandelt werden.

6.4.5 Aktive Regeln zur Prioritätenvergabe

Bei der Prioritätenvergabe handelt es sich um eine zentrale Tätigkeit des Prioritätenvergabe- und -optimierungssystems. Die planmäßige Vergabe geschieht auf Anfrage oder Anforderung des Dispositionssystems für den entsprechenden Betriebstag, wenn es die entsprechenden Angaben benötigt. Eine (unverbindliche) Anfrage kann prinzipiell zu jedem Zeitpunkt erfolgen. Die (verbindliche) Anforderung der Abschnittsweisen Zugprioritäten erfolgt unmittelbar vor Einfahrt des betroffenen Zuges in den Prioritätenabschnitt sowie für Anschlußprioritäten bei Ankunft des aufnehmenden Zuges in einen Bahnhof mit Anschlußverbindungen. Alternativ besteht die Möglichkeit, die zu verwendenden Standard- und Störfallprioritäten in einer für das Dispositionssystem und Prioritätenvergabe- und -optimierungssystem gleichzeitig auszulesenden Datenbank abzulegen und den Prozeß der planmäßigen Prioritätenvergabe vollständig dem Dispositionssystem zu überlassen.

Bei einer endgültigen Prioritätenanfrage oder -anforderung wird anhand des maßgeblichen Speziellen Betriebszustandes die zutreffende abgeleitete Standard- oder Störfallpriorität übermittelt und die Festschreibung der Priorität protokolliert. Zusätzlich wird im Falle der endgültigen Anforderung der maßgebliche Spezielle Betriebszustand für zukünftige Auswertungen im Rahmen der Prioritätenoptimierung gespeichert. Daraus ergeben sich die folgenden relativ einfachen aktiven Regeln (hier exemplarisch für Abschnittsweise Zugprioritäten):

```
on anfrage_abs_zugpr(Mz, Pabs, Bt)
if abs_zugpr(Mz, Pabs, Bt, Rg)
do signal_abs_zugpr(Mz, Pabs, Bt, Rg).

on anford_abs_zugpr(Mz, Pabs, Bt)
if abs_zugpr(Mz, Pabs, Bt, Rg)
do {
      signal_abs_zugpr(Mz, Pabs, Bt, Rg);
      + zugpr_aktiv(Mz, Pabs, Bt)
   }.

on anford_abs_zugpr(Mz, Pabs, Bt)
if maßg_sp_bz_zug(Sp_Bz, Mz, Pabs, Bt)
do + maßg_sp_bz_zug_verg(Sp_Bz, Mz, Pabs, Bt).
```

Nach erfolgter Durchfahrt durch den Prioritätenabschnitt wird das entsprechende zugpr_aktiv(…)-Fakt für den Modellzug wieder gelöscht (hier ohne Regeldarstellung).

Beim plötzlichen Auftreten schwererer Störungen können störfallbedingte Prioritätenvergaben „unter rollendem Rad“ durchgeführt werden. Dies betrifft primär Abschnittsweise Zugprioritäten während der Zug bereits auf dem entsprechenden Prioritätenabschnitt verkehrt. Für Anschlußprioritäten ist es theoretisch auch möglich, daß eine Anschlußpriorität während der Wartezeit des aufnehmenden Zuges noch einmal angepaßt wird. Dies spielt aber wegen des in der Regel relativ kurzen Zeitraums zwischen Prioritätenvergabe und der Anschlußaufnahme praktisch keine so große Bedeutung und soll daher hier nicht detailliert behandelt werden.

Grundsätzlich sollten im Sinne einer relativ stetigen Betriebsdurchführung Prioritätenwechsel von Abschnittsweisen Zugprioritäten möglichst nur beim Aufenthalt bzw. Durchfahrt in den „Grenzbahnhöfen“ zwischen zwei Prioritätenabschnitten erfolgen. Hier bestehen laut Anforderungen an diese Bahnhöfe ausreichende Pufferkapazitäten zur Verfügung, um einen Prioritätenwechsel und evtl. damit verbundene längere Wartezeiten ausreichend implementieren zu können. Auf den Prioritätenabschnitten selber ist es nicht sicherzustellen, daß eine nachträgliche Anpassung der Priorität auch wirklich vom Dispositionssystem so umgesetzt werden kann. Unter Umständen sind nur Konfliktlösungen möglich, die der Vorgabe des Prioritätenvergabe- und -optimierungssystems entgegengesetzt sind.

Störfallbedingte Prioritätenvergaben sollten daher möglichst selten und nur in wirklich wichtigen Fällen durchgeführt werden. Daher empfiehlt sich die Einschränkung auf Initialisierungen/Aktualisierungen/Terminierungen von Speziellen Betriebszuständen mit einem hohen Gewicht. In allen sonstigen Änderungsfällen maßgeblicher Spezieller Betriebszustände sollen störfallbedingte Prioritätenvergaben ausgeschlossen sein.

Ferner ist zu beachten, daß es im Falle von störfallbedingten Prioritätenvergaben keinen eindeutigen maßgeblichen Speziellen Betriebszustand des Zuges auf dem Prioritätenabschnitt gibt. Daher sollten die entsprechenden Betriebstage nicht als Referenzwerte für die Toleranzgrenzenüberschreitungen und Prioritätenoptimierung verwendet werden. Dies kann dadurch sichergestellt werden, daß die jeweiligen bei der planmäßigen Prioritätenvergabe gespeicherten Werte in diesem Falle ersatzlos gelöscht werden. (Zur Erinnerung: Bei Verwendung von Standardprioritäten wird der Wert NULL vergeben.)

Für eine störfallbedingte Prioritätenänderung muß eine explizite Freigabe seitens des Benutzers gegeben werden, was im folgenden syntaktisch durch die Angabe „with commit“ am Ende der Regel ausgedrückt wird. Daraus ergeben sich folgende aktive Regeln für die störfallbedingte Prioritätenvergabe:

```
on + maßg_sp_bz_zug(Sp_Bz, Mz, Pabs, Bt)
if zugpr_aktiv(Mz, Pabs, Bt),
  Sp_Bz.Typ.Gewicht >= [Schwellenwert],
  abs_zugpr(Bt,Mz,Pabs,Rg)
do {
     signal_abs_zugpr(Mz, Pabs, Bt, Rg);
     – maßg_sp_bz_zug_verg(_, Mz, Pabs, Bt)
   }
with commit.

on – maßg_sp_bz_zug(Sp_Bz, Mz, Pabs, Bt)
if zugpr_aktiv(Mz, Pabs, Bt),
  Sp_Bz.Typ.Gewicht >= [Schwellenwert],
  maßg_sp_bz_zug(Sp_Bz_Neu, Mz, Pabs, Bt),
  Sp_Bz_Neu.Typ.Gewicht < [Schwellenwert],
  abs_zugpr(Bt,Mz,Pabs,Rg)
do {
     signal_abs_zugpr(Mz, Pabs, Bt, Rg);
     – maßg_sp_bz_zug_verg(_, Mz, Pabs, Bt)
   }
with commit.
```

6.4.6 Aktive Regeln zur Prioritätenoptimierung

Die Prioritätenoptimierung ist neben der Prioritätenvergabe der bedeutendste Prozeß des Systems. Hier geht es darum, aufgrund von Toleranzgrenzenüberschreitungen für verursachte Bonus-Malus-Zahlungen Anpassungen der vorgegebenen Standard- und Störfallprioritäten vorzunehmen.

Das Optimierungssystem versucht, gemäß des in 6.3.2 angegebenen generellen Algorithmus entsprechende Prioritätenänderungen vorzunehmen. Dabei muß einerseits darauf geachtet werden, daß das Ergebnis legal bzgl. der normativen Regeln ist, und andererseits, daß eine Änderung von Prioritäten nicht zu einer bereits früher verwendeten lokalen Prioritätenkonstellation führt, die bereits als ungünstig bewertet wurde.

Es gibt dementsprechend drei alternative Ergebnisse nach der Detektierung einer Toleranzgrenzenüberschreitung:

- Absenkung oder Erhöhung einer Priorität entsprechend der Toleranzgrenzenüberschreitung
- Rücksetzung der Priorität auf einen nach derzeitigem Kenntnisstand optimalen Wert
- Beibehaltung des Status Quo

In diesem Abschnitt steht die Formulierung der entsprechenden aktiven Regeln im Mittelpunkt.

Im Falle einer Toleranzgrenzenüberschreitung für verursachte Bonus-Malus-Zahlungen durch einen relativ hochrangigen Zug in einem Prioritätenabschnitt muß eine Absenkung der Abschnittsweisen Zugpriorität dieses Zuges auf dem entsprechenden Prioritätenabschnitt überprüft werden. Der neue Wert kann von der deduktiven Komponente bestimmt werden. Details zur Ermittlung wurden in 5.2 genannt. Wenn diese Änderung nicht den normativen Regeln widerspricht und keine bereits als ungünstig eingestufte Konstellation erreicht wird, kann die Änderung erfolgen und entsprechende Reset-Ereignisse für den Zug und in bzgl. Dispositionsmaßnahmen mit diesem Zug in möglichem Konflikt stehende Züge mit den entsprechenden Speziellen Betriebszuständen getriggert werden.

Ansonsten soll bei der Existenz einer bisher optimalen Prioritätenkonstellation diese wiederhergestellt werden und der Reset erfolgen bzw. (bei Nichtvorhandensein) eine Nachricht über die nicht mögliche Optimierung an die Benutzerschnittstelle weitergeleitet werden.

Es ist zu beachten, daß das System in der Weise konstruiert wird, daß einerseits keine mehrfache Reaktion auf gleichzeitige Toleranzgrenzenüberschreitungen desselben Zuges bzw. bzgl. Dispositionsmaßnahmen möglicherweise mit diesem in Konflikt stehende Züge erfolgen. Daher wird immer nur auf die von der deduktiven Komponente ermittelte gewichtigste Toleranzgrenzenüberschreitung für eine Bonus-Malus-Zahlung auf einem Prioritätenabschnitt reagiert. Der Reset-Mechanismus sorgt dafür, daß die folgende, von der deduktiven Komponente signalisierte gewichtigste Toleranzüberschreitung nicht in einem Verhältnis

zur gerade korrigierten Priorität steht. Alle Prioritätenoptimierungen treten erst verzögert zum nächsten Betriebstag in Kraft.

Hierfür ergibt sich folgende Beispielregel:

```
on + gew_tg_überschr(BMZ, Pabs, Mz, hochr, Sp_Bz_Typ, Tg)

if abgesenkte_zugpr(Pabs, Mz, Sp_Bz_Typ, NeuPr),
   not zugpr_ungünstig(Pabs, Mz, Sp_Bz_Typ, NeuPr),
   reset_zuege(Pabs, Mz, Sp_Bz_Typ, Reset_Mz, Reset_Sp_Bz_Typ)

do
{
    + prior_reset(Pabs, Reset_Mz, BMZ.Bt, Reset_Sp_Bz_Typ);
    + vorm_pr_absenkung (Mz, Pabs, Sp_Bz_Typ, NeuPr)
}
with commit.
```

Man beachte, daß das Fakt

```
abgesenkte_zugpriorität(…).
```

durch die normative Komponente überwacht werden kann, so daß im Rahmen der aktiven Regel keine weitere explizite Prüfung bzgl. Legalität der neuen Priorität erfolgen muß.

Ferner erfolgt die Prioritätenabsenkung nur als „Vormerkung“ und nicht als explizite Datenbankänderung. Die eigentliche Änderung kann dann über eine weitere aktive Regel von Typ „deferred“ erfolgen, die zum Wechsel des Betriebstags „gefeuert“ wird. Hierbei wird im Falle der Änderung einer Störfallpriorität sichergestellt, daß die Änderung für alle mit der entsprechenden Störfallpriorität verbundenen Speziellen Betriebszustandstypen durchgeführt wird. Zur Durchführung des Aktionsteils ist ferner die Freigabe (commit) des Benutzers erforderlich.

Ferner ist die folgende Regel möglich:

```
on + gew_tg_überschr (BMZ, Pabs, Mz, hochr, Sp_Bz_Typ, Tg)

if not abgesenkte_zugpr(Pabs, Mz, Sp_Bz_Typ, NeuPr),
   bisher_optimale_lsg(Pabs, Mz, Sp_Bz_Typ, NeuPr),
   reset_zuege(Pabs, Mz, Sp_Bz_Typ, Reset_Mz, Reset_Sp_Bz_Typ)

do
{
    + prior_reset(Pabs, Reset_Mz, BMZ.Bt, Reset_Sp_Bz_Typ);
    + vorm_pr_absenkung (Mz, Pabs, Sp_Bz_Typ, NeuPr)
}
with commit.
```

Hier wird im Falle von nicht vorhandenen legalen Optimierungsmöglichkeiten die unter gegebener Konstellation bisher sich als am besten bewährte Priorität gewählt.

Eine weitere Regel (die hier nicht mehr im Detail angegeben werden soll) besagt, daß falls sich die vorgeschlagene Optimierung bereits als ungünstig erwiesen hat, ebenfalls die sich bisher am besten bewährte Priorität gewählt werden soll. Schließlich soll in allen übrigen Fallen (kein „legaler“ Vorschlag vorhanden und für die entsprechende Konstellation kein

bisheriges Optimum vorhanden) der Status Quo beibehalten werden. Analog kann auf Toleranzgrenzenüberschreitungen durch relativ niederrangige Züge reagiert werden, wobei hier eine relative Anhebung der Abschnittsweisen Zugpriorität um ein, zwei oder drei relative Rangstufen erfolgen muß.

Für Anschlüsse wird in entsprechender Weise vorgegangen. Hier erfolgt eine Erhöhung der Wartezeiten, wenn die Toleranzgrenze für verursachte Bonus-Malus-Zahlungen durch gebrochene Anschlüsse an diesem Anschluß überschritten wurde. Eine Absenkung erfolgt, falls durch Einhalten des Anschlusses eine Toleranzgrenze für die Verursachung (im weiteren Fahrtverlauf entstandener) prioritätenbedingter Bonus-Malus-Zahlungen verletzt wird.

Eine vollständige Darstellung sämtlicher im Zusammenhang des im Rahmen dieser Arbeit entwickelten Prioritätenvergabe- und -optimierungssystems verwendbaren Regelkonzepte bzw. Regeln würde den Rahmen dieser Arbeit bei weitem sprengen. Daher konnten an dieser Stelle lediglich eine Übersicht zu den wichtigsten möglichen Einsatzgebieten sowie einige wenige exemplarische Regeln vorgestellt werden, die stellvertretend das Grundprinzip verdeutlichen sollen.

Das vorgestellte Prioritätenvergabe- und -optimierungssystem kann statt mit den angegebenen Regelkonzepten alternativ auch selbstverständlich mit Hilfe einer objektorientierten, imperativen oder deklarativen Programmiersprache implementiert werden. Die Vorteile des regelbasierten Ansatzes wurden im Rahmen dieser Arbeit ausführlich dargestellt.

7 Prototypische Implementierung, Fazit und Perspektiven

Basierend auf dem im vorausgehenden Kapiteln entwickelten Systementwurf zur Prioritätenvergabe und -optimierung wurde seitens des Verfassers zur Prüfung der praktischen Realisierbarkeit dieses Entwurfs und zur Gewinnung weiterer Erfahrungen eine Implementierung eines prototypischen Prioritätenvergabe- und -optimierungssystems für Demonstrations- und Testzwecke namens POS vorgenommen.

Im folgenden Kapitel soll daher nach einem einführenden Überblick auf POS zunächst ein Einblick in interessante Aspekte der Implementierung erfolgen. Schließlich soll die Arbeit mit einem zusammenfassenden Fazit und einem Ausblick auf weitere auf die Arbeit aufbauende Forschungsaktivitäten abgeschlossen werden.

7.1 Prototypische Implementierung von POS

Das Prioritätenvergabe- und -optimierungssystem POS wurde als eine Microsoft Access Datenbankanwendung implementiert. Eigentlich handelt es sich bei Access nicht unbedingt für eine optimale Wahl für ein regelbasiertes System, da dieses DBMS keine generell eingebaute Unterstützung für die benötigten aktiven Regeln/Trigger bietet. Das Verhalten der Trigger kann in Access allerdings mittels Visual Basic-Modulen in imperativer Programmierform „simuliert“ werden. Aus Gründen vereinfachender Systemwartbarkeit und wegen besserer Leistungsfähigkeit empfiehlt es sich daher, ein im wirklichen Einsatz befindliches Prioritätenvergabe- und -optimierungssystem grundsätzlich auf einem DBMS mit derartiger Triggerunterstützung zu implementieren. Für die bei dieser Implementierung primären Demonstrations- und Testzwecke wurde dennoch auf Access zurückgegriffen, da hier so ein einfacher Einsatz an einem PC ohne nennenswerten Installationsaufwand sichergestellt werden kann.

Da es sich um ein prototypisches Demonstrationssystem handelt, beinhaltet POS neben der Prioritätenvergabe- und -optimierungskomponente auch die zu diesem Einsatz benötigten (eigentlich „externen“) Dispositions- und Qualitätsüberwachungskomponenten. Dabei wurden diese unterschiedlichen Komponenten allerdings voneinander getrennt implementiert um der eigentlich funktionale Trennung dieser Systeme besser zu entsprechen. Ferner erfolgen die Implementierungen der „externen“ Dispositions- und Qualitätsüberwachungskomponenten mittels imperativer Programmierung mit einer Datenverwaltung im Hauptspeicher, wohingegen die Prioritätenvergabe- und -optimierungskomponente weitgehend (mit Ausnahme der bereits erwähnten „simulierten“ Trigger) unmittelbar durch das regelbasierte DBMS selber erfolgt.

Selbstverständlich besteht für POS keine Möglichkeit, direkt auf Betriebsdaten eines Eisenbahninfrastrukturbetreibers zurückzugreifen. Dementsprechend muß der eigentliche Eisenbahnbetrieb selber ebenfalls durch POS simuliert werden. Für Test- und Demonstrationszwecke der Prioritätenoptimierung ist eine relativ „grobe“ Betriebssimulation völlig ausreichend. Dazu erzeugt POS mittels eines Pseudozufallsprozesses entsprechend vorgegebener Verteilungen Urverspätungen sowohl für Haltezeitverzögerungen wie auch für

verlangsamte Fahrten auf einem Streckenabschnitt. Es wird ein planmäßiger Fahrtzeitenzuschlag von 5% angenommen, um das „Aufholen“ kleinerer Verspätungen zu ermöglichen. Belegungskonflikte auf Streckenabschnitten werden vereinfachend durch Auswertungen von Planankunfts- bzw. -abfahrtszeiten gegenüber den „realen“ Zeiten der Züge auf dem Streckenabschnitt bzgl. unplanmäßiger Überholungen detektiert. Die Lösung der Belegungskonflikte erfolgt aufgrund der Prioritäten und der Geschwindigkeitsdifferenzen. Niederrangige Züge werden um 5 Minuten verzögert, falls der höherrangige Zug relativ schneller ist (einfache Überholung). Sie erreichen den Zielbahnhof des Streckenabschnitts 2 Minuten nach dem hochrangigen Zug, falls dieser relativ langsamer ist (hinterherfahren). Anschlußkonflikte werden entsprechend der Anschlußprioritäten (vorgegebene maximale Wartezeit auf Anschlußzubringer) entschieden. Das Auslesen der korrekten Standard- oder Störfallpriorität aus der POS-Datenbank erfolgt für Zugprioritäten bei tatsächlicher Einfahrt des Zuges in den Streckenabschnitt und für Anschlußprioritäten bei Erreichen der Abfahrtszeit.

Aufgrund der beschriebenen Betriebssimulation wird seitens POS ein Dispositionsfahrplan generiert, der dann mittels der Qualitätsüberwachungskomponente analysiert wird. Hierbei wird geprüft, ob wegen Verspätungen an vorgegebenen Meßpunkten oder Anschlußbrüchen Bonus-Malus-Zahlungen ausgelöst werden. Falls dies der Fall ist, erfolgt entsprechend des in dieser Arbeit erläuterten Verfahrens eine für die Prioritätenoptimierungskomponente wichtige Aufschlüsselung von Bonus-Malus-Anteilen auf die eigentlichen Verursacher. Dies erfolgt mit Hilfe einer Analyse von Urverspätungen und Dispositionsfahrplänen der betroffenen Züge. Die Ergebnisse des Qualitätsüberwachungsprozesses werden ebenso wie die Ergebnisse des Dispositionsprozesses in die POS-Datenbank geschrieben.

Zum Ablauf eines Betriebstages (mittels „simulierter“ deferred Trigger) wird der eigentliche POS-Optimierungsprozeß durchgeführt. Hierbei wird aufgrund vom Benutzer für jeden Störfall und eine vorgegebene Anzahl von n Betriebstagen vorgegebener Toleranzgrenzen für Bonus-Malus-Verursachungsanteile auf jeden Streckenabschnittskilometer eines Zuges bzw. durch jeden Anschluß (getrennt ausgewertet für Einhaltung bzw. Bruch) auf Basis des Qualitätsüberwachungsoutput die Überschreitungen von Toleranzgrenzen detektiert. Dazu müssen die letzten n Betriebstage ermittelt werden, bei denen die aktuell verwendete Standard- oder Störfallpriorität unverändert vergeben wurde. Die Toleranzgrenzenüberschreitungen werden absteigend anhand des durchschnittlich pro Tag verursachten Bonus-Malus-Betrages absteigend sortiert. Dies erfolgt alles mittels deduktiver Regeln.

„Simulierte Trigger“ reagieren nun direkt auf die jeweils „gravierendste“ Toleranzgrenzenüberschreitung. Bei Überschreitung für einen Anschluß erfolgt entsprechend eine Erhöhung (wg. Bruch) oder Erniedrigung (wg. Einhaltung) der Anschlußpriorität um 10%. Bei Überschreitung für einen Streckenabschnitt erfolgt eine Aufwertung (wg. Bonus-Malus als niederrangiger Zug) oder Abwertung (wg. Bonus-Malus als hochrangiger Zug) um eine Rangstufe. Ferner erfolgt gleichzeitig eine Bestätigung (Update) der Standardzugprioritäten von Zügen in der „Umgebung“ des geänderten Zuges. Diese Menge wird wiederum durch eine deduktiv aus vergangenen Konfliktlösungsdaten abgeleitet. Diese Bestätigung dient dazu, die Züge in dieser Umgebung aus den abgeleiteten Toleranzgrenzenüberschreitungen zu eliminieren, um parallele sich gegenseitig beeinflussende Prioritätenänderungen zu vermeiden. Schließ-

lich erfolgt nach einem Zugprioritätenoptimierungsschritt mittels deduktiver Regeln eine Verifikation, ob ein ermittelter neuer Prioritätenvektor sich bereits im Betrieb als ungünstig erwiesen hat. In diesem Fall kann ein sich als bisher optimal erwiesener Vektor als Alternativvektor vorgeschlagen werden.

Nach dem Optimierungsprozeß werden die neuen aktuellen Standard- und Störfallprioritäten in der POS-Datenbank abgelegt und stehen der Dispositionskomponente bei Bedarf zum Abruf bereit.

In Testläufen zeigt sich durch den Einsatz von POS ein Optimierungspotential bzgl. verursachter Bonus-Malus-Zahlungen abhängig von den pseudozufällig erzeugten Urverspätungen zwischen 10 und 20 % gegenüber einem Lauf mit ausgeschalteter Prioritätenoptimierungskomponente. Allgemeine Aussagen zur Leistungsfähigkeit unabhängig von einem konkreten Anwendungsfall sind allerdings kaum möglich, da diese neben den pseudozufällig generierten Urverspätungen stark vom hinterlegten Fahrplan und der Qualität der initialen Standardprioritäten abhängt.

7.2 Fazit und Perspektiven

In dieser Arbeit wurde die Entwicklung eines regelbasierten Systems zur Betriebsdisposition mit regelbasiert optimierten Prioritäten im Eisenbahnverkehr beschrieben. Das entworfene Prioritätenvergabe- und -optimierungssystem dient einem angeschlossenen Dispositionssystem als Entscheidungsunterstützung im Falle von auftretenden Konflikten und arbeitet losgelöst vom eigentlichen Konfliktlösungsprozeß. Im Gegensatz zu den heute verwendeten Dispositionssystemen arbeitet das System mit dynamischen anstelle von statischen Abschnittsweisen Zug- und Anschlußprioritäten, auf deren Basis das angeschlossene Dispositionssystem selbständig mögliche konkrete Konfliktlösungen ermitteln und praktisch umsetzen kann. Dabei wird zwischen regulärem Betrieb und speziellen, bzgl. Prioritäten besonderer Behandlung bedürfender Störfallszenarien unterschieden. Mit Hilfe eines angegliederten Qualitätsüberwachungssystems kann eine Bewertung des disponierten Betriebs auf Basis eines der aktuellen europäischen Gesetzgebung entsprechenden Bonus-Malus-Systems erfolgen. Diese dient als Grundlage einer weiteren Optimierung der Prioritäten. Entsprechend wird im hier entwickelten System eine zweistufige Optimierung vorgenommen. Es erfolgt eine kurzfristige und schnelle Ad-hoc-Prioritätenvergabe von abgespeicherten Standard- bzw. Störfallprioritäten, die wiederum mittel- bis langfristig in einem separaten Prioritätenoptimierungsprozeß ermittelt werden.

Für den Kernprozeß der Prioritätenverwaltung und -optimierung kann auf vorhandene deduktive, aktive und normative Komponenten einer Datenbankverwaltung zurückgegriffen werden.

Der dargestellte Ansatz verfolgt das Ziel, Vorteile aus den Optimierungs- und heuristischen Verfahren für die Eisenbahnbetriebsdisposition in einem Gesamtkonzept zu vereinen.

Ferner wurde im Rahmen dieser Arbeit ein prototypische Implementierung des Prioritätenvergabe- und -optimierungssystems POS vorgenommen, womit die Fähigkeit zur Umsetzung mit auf dem Markt vorhandenen Datenbankmanagementsystemen auch praktisch gezeigt wurde.

Selbstverständlich kann im Rahmen einer solchen Arbeit bei der Entwicklung eines derart komplexen Systems nicht jedes Detail und jeder pathologische Fall ausführlich dargestellt werden, dennoch sollten die wesentlichen Prinzipien deutlich gemacht worden sein, so daß sich weiterführende Aspekte leicht aus diesen ableiten lassen.

Bzgl. weiterer auf diese Arbeit aufbauenden wissenschaftlichen Forschungsarbeiten sind als Folgeschritte abschließend die folgenden Punkte zu empfehlen

- Vertiefte Forschung im Bereich Bonus-Malus-Systeme bzgl. deren Abbildung realer verkehrswirtschaftlicher Zusammenhänge. Letztendlich kann die Qualität des vom Prioritätenvergabe- und -optimierungssystem gelieferten Outputs nur verwertbar sein, wenn im System ein geeignetes Bonus-Malus-System abgelegt ist,
- Implementierung der in dieser Arbeit aufgezeigten und prototypisch implementierten Grundlagen in einem Server-basierten Datenbanksystem mit vollständiger interner Trigger-Unterstützung,
- Integration in ein kommerziell genutztes gleis- und signalgenaues Dispositionssystem,
- konkrete kundenspezifische Anpassung des in dieser Arbeit aufgezeigten allgemeinen Systementwurfs auf die Bedürfnisse eines konkreten Eisenbahninfrastrukturbetreibers,
- Weiterentwicklung des Optimierungsansatzes für weitere prioritätenbasierte Prozesse im Verkehrsbereich und ggf. über den Verkehrsbereich hinaus.

Mögliche Anwendungsgebiete für ein ähnliches System sind in der Regel da gegeben, wo verschiedene Prozesse unterschiedlicher Priorität um beschränkte Betriebsmittel konkurrieren müssen und für verzögerte Abschlüsse der Prozesse ein entsprechendes Pönal anfällt. Im Eisenbahnbereich wäre beispielsweise die Echtzeitzuweisung von betriebsbereiten Lokomotiven auf Züge ein ähnliches Einsatzgebiet. Auch die Priorisierung von Rangiervorgängen im Rangierbahnhof wäre hier denkbar. Im weiteren Verkehrsbereich ist ein Einsatz z. B. im Zusammenhang mit Schiffsschleusungen bzw. mit Start- und Landevorgängen im Luftverkehr möglich. Im Bereich Logistik kann eine Prioritätenvergabe für unterschiedliche Sendungen in einem Netzwerk mit beschränkter Kapazität erfolgen. Außerhalb des Verkehrsbereiches sind Einsatzgebiete z. B. im Zusammenhang mit Auftragsverwaltungen, kurzfristigen Dienstplanerstellungen oder bei der Prozeßverwaltung in Betriebssystemen denkbar.

Abkürzungsverzeichnis

Verwendete Abkürzungen für Regelvariablen in Kapitel 6.4:

Bt: Betriebstag

Mz: Modellzug

Pabs: Prioritätenabschnitt

Rg: Rangstufe

RzGz: Reisezug oder Güterzug

Anschl: Anschluß

Max_Wz: maximale Wartezeit

Sp_Bz: Spezieller Betriebszustand

Sp_Bz_Typ: Spezieller Betriebszustandstyp

BMZ: Bonus-Malus-Zahlung;

NrHr: Niederrangig oder Hochrangig

Pabw: Planabweichung

Vhalt: Verkehrshalt;

Pos: Position im Zuglauf

Relev_Bt: Relevanter Betriebstag

Reset-Bt: Reset-Betriebstag

Tg: Toleranzgrenze

VglBt: Vergleichsbetriebstag

BM_Betr Bonus-Malus-Betrag

Tages_BM Tages-Bonus-Malus-Betrag

Vgl_BM Bonus-Malus-Betrag in Vergleichsbetriebstagen

Vh: Verkehrshalt

ProReal Prognostiziert oder Real

Gew Gewicht

Zt Zeitstempel

Sp_Bz_Vg Spezieller Betriebszustand bei Prioritätenvergabe

NeuPr Neue Priorität

Reset_Mz Mit einem Modellzug möglicherweise in Konflikt stehender Modellzug

Reset_Sp_Bz_Typ Spezieller Betriebszustandstyp für mit einem Modellzug möglicherweise in Konflikt stehenden Modellzug

Abbildungsverzeichnis

Literaturverzeichnis

[Ack98] **T. Ackermann**, Die Bewertung der Pünktlichkeit als Qualitätsparameter im Schienenverkehr auf der Basis der direkten Nutzermessung
Dissertation, Universität Stuttgart, 1999

[AEG93] Allgemeines Eisenbahngesetz (AEG) vom 27.12.1993
BGBl. I S. 2378, 2396, 1994 I S. 2439
zuletzt geändert am 29.7.2009 (BGBl. I S. 2542)

[Bär96] **M. Bär**, Konflikte im Eisenbahnbetrieb – Versuch einer Systematisierung, Schriftenreihe des Instituts für Verkehrssystemtheorie und Bahnverkehr
TU Dresden, Band 2, 1996, S. 138-149

[Bär98] **M. Bär**, Planungs- und Dispositionssysteme für den Bahnbetrieb
Eisenbahningenieur (49) 3/1998, S. 19-26

[Bar01] **S. Barke, H. Heike**, Leitsysteme von Siemens für Betriebszentralen
Signal + Draht 93, 4/2001, S. 5-11

[Bay80] **R. Beyersdorf**, Die Rechnerunterstütze Zugüberwachung
Eisenbahntechnische Rundschau, 29 (1980), Heft 11, S. 769-772

[Boo91] **G. Booch**, Object-Oriented Design with Applications
Benjamin/Cummings, 1994

[Boo99] **G. Booch, J. Rumbaugh, I. Jacobson**, Das UML-Benutzerhandbuch
Addison-Wesley, 1999

[Bre03] **S. Breu, G. Thiemt, M. Kant**, LeiDis-N – Netzdisposition in der Netzleitzentrale der DB AG in Frankfurt
Signal + Draht (95), 6/2003, S. 14-19

[Bro85] **L. Brownston, R. Farrell, E. Kant**, Programming Expert Systems in Ops5: An Introduction to Rule-Based Programming
Addison-Wesley, 1985

[Brü95] **O. Brünger**, Konzeption einer Rechnerunterstützung für die Feinkonstruktion von Eisenbahnfahrplänen
Dissertation, Veröffentlichungen des Verkehrswissenschaftlichen Instituts der RWTH Aachen, Heft 51, 1995

[Cha74] **D. Chamberlin, R. Boyce**, SEQUEL: A Structured English Query Language
SIGMOD Workshop, Vol. 1 1974: 249-264

[Che76] **P. Chen**, The Entity-Relationship Model – Toward a Unified View of Data
In: ACM Transactions on Database Systems 1/1/1976 ACM-Press

[Chr04] **L. Christiany**, Übergangszeiten als neue Grundlage für Anschlüsse im Personenverkehr im Fahrplan 2004/2005
BahnPraxis 12/2004, S. 138-139

[Clo94] **W. Clocksin, C. Mellish**, Programming in Prolog
Springer, Forth Edition, 1994

[Cor01] **T. Cormen et al.**, Introduction to Algorithms
MIT Press and McGraw-Hill, Second Edition, 2001

[Dan85] **H. Dannenberg**, Ordered-Search-Algorithmus zur rechnergestützten Disposition von Eisenbahnbetriebsabläufen
Dissertation, Wissenschaftliche Arbeit Nr. 54 des Instituts für Verkehrswesen, Eisenbahnbau und -betrieb der Universität Hannover, 1985

[Dat95] **C. Date**, An introduction to database systems
Addison-Wesley, 6th edition, 1995

[DBNe09] **DB Netz AG**, Das Anreizsystem Trasse der DB Netz AG, Maßnahmen zur Verringerung von Störungen und Erhöhung der Leistungsfähigkeit auf dem Netz 2009

[DBR408] **DB Netz AG**, Richtlinie 408, Züge fahren und Rangieren

[DBR420] **DB Netz AG**, Richtlinie 420, Betriebszentralen DB Netz AG

[Din06] **M. Dincbas,** Constraint programming — an alternative to expert systems
Lecture Notes in Computer Science, Springer, 2006

[Dit00] **K. Dittrich, S. Gatziu**, Aktive Datenbanksysteme, Konzepte und Mechanismen
dpunkt.verlag, 2000

[EBO67] Eisenbahn-Bau- und Betriebsordnung vom 8.5.1967
BGBl. II S. 1563
zuletzt geändert am 19.3.2008 (BGBl. I S. 467)

[EIBV05] Verordnung über den diskriminierungsfreien Zugang zur Eisenbahninfrastruktur und über die Grundsätze zur Erhebung von Entgelt für die Benutzung der Eisenbahninfrastruktur
(Eisenbahninfrastruktur-Benutzungsverordnung – EIBV) vom 3.7.2005
BGBl. I S. 1566
zuletzt geändert am 3.6.2009 (BGBl. I S. 1235)

[Eif06] **A. Eifer**, Objektorientierte, regelbasierte Modellierung komplexer Szenarien im Straßenverkehr
Diplomarbeit an der Rheinischen Friedrich-Wilhelms-Universität Bonn, 2006

[EG01] Richtlinie 2001/14/EG des Europäischen Parlaments und des Rates vom 26. Februar 2001 über die Zuweisung von Fahrwegkapazität der Eisenbahn und die Erhebung von Entgelten für die Nutzung der Eisenbahninfrastruktur
ABl. L 75 vom 15.3.2001, S.29

[EG07] Verordnung (EG) Nr. 1371/2007 des Europäischen Parlaments und des Rates vom 23. Oktober 2007 über die Rechte und Pflichten der Fahrgäste im Eisenbahnverkehr
ABl. L 315 vom 3.12.2007, S. 14

[EG08] **Kommission der Europäischen Gemeinschaften**, Vorschlag für eine Verordnung des Europäischen Parlaments und des Rates zur Schaffung eines europäischen Schienennetzes für einen wettbewerbsfähigen Güterverkehr
SEK(2008) 3028, SEK(2008) 3029), 2008

[EuPa05] **European Patent Office**, Method for resolving Conflicts in a trackbound transportation system
EP 1 500567 B1, 2005

[EVO38] Eisenbahnverkehrsordnung (EVO) vom 8.9.1938
RGBl. II S. 663
in der Neufassung vom 20.4.1999 (BGBl. I 782)
zuletzt geändert am 26.5.2009 (BGBl. I S.1146)

[EWG91] Richtlinie 91/440/EWG des Rates vom 29. Juli 1991 zur Entwicklung der Eisenbahnunternehmen der Gemeinschaft
Amtsblatt Nr. L 237 vom 24.08.1991 S. 0025 – 0028

[Fai06] **T. Faison**, Event-Based Programming: Taking Events to the Limit
Apres, 2006

[Fay99] **A. Fay**, Wissensbasierte Entscheidungsunterstützung für die Disposition im Schienenverkehr – eine Anwendung von Fuzzy-Petrinetzen
Dissertation, Fortschritt-Berichte VDI, Reihe 12, Nr. 386, 1999

[Fay00] **A. Fay**, Dezentrale Steuerung des Schienenverkehrs durch autonome Agenten
Signal+Draht 92, Heft 3, 14-18, 2000

[Flo62] **R. Floyd**, Algorithm 97: Shortest Path
Communications of the ACM 5 (6), S. 345, 1962

[Freu87] **H. Freudenberger, G. Gottfried, I. Schmidt**, XPS 1 – der erste Prototyp eines Expertensystems bei der DB
Signal + Draht 79 1/2, 1987

[Gal84] **H. Gallaire, J. Minkler, J. Nicholas**, Logic and Databases: A Deductive Approach
ACM Computing Surveys, Juni 1984, 16(2), S. 153-185

[Gan80] **R. Gandy**, Church's Thesis and Principles for Mechanisms
In Barwise, J., Keisler, H.J., Kunen, K. (eds) 1980. The Kleene Symposium, Amsterdam, North-Holland.

[GG49] Grundgesetz für die Bundesrepublik Deutschland vom 23.5.1949
BGBl. I S. 1
zuletzt geändert am 29.7.2009 (BGBl. I S. 2248)

[Gil62] **A. Gill**, Introduction to the Theory of Finite-state Machines
McGraw-Hill, New York 1962

[Glo97] **F. Glover, M. Laguna**, Tabu Search
Kluwer, 1997

[Gro66] **A. Groll, H. Kayser**, Fahrplanwesen
Josef Keller Verlag Starnberg, 1. Aufl., 1966

[Grö02] **T. Gröger**, Simulation der Fahrplanerstellung auf der Basis eines hierarchischen Trassenmanagements und Nachweis der Stabilität der Betriebsabwicklung
Dissertation, Veröffentlichungen des Verkehrswissenschaftlichen Institutes der RWTH Aachen, Heft 60, 2002

[Hap59] **O. Happel**, Sperrzeiten als Grundlage für die Fahrplankonstruktion, Eisenbahntechnische Rundschau 8 (1959) H. 2, S. 79-90

[Hau99] **D. Hauptmann, R. Kaminsky, A. Radtke**, Integrierte Planung des Eisenbahnbetriebs mit SIMU++
Signal + Draht 91 7/8, 1999

[Hau00] **D. Hauptmann**, Automatische und diskriminierungsfreie Ermittlung von Fahrplantrassen in beliebig großen Netzen spurgeführter Verkehrssysteme
Dissertation, Wissenschaftliche Arbeit Nr. 54 des Instituts für Verkehrswesen, Eisenbahnbau und –betrieb der Universität Hannover, 2000

[Hel04] **S. Heller, T. Schaer**, DisKon – Dispositions- und Konfliklösungsmanagement der DB AG
Eisenbahningenieur (55), 9/2004, S. 103-122

[Hoa69] **C.A.R. Hoare**, An axiomatic basis for computer programming
Communications of the ACM, 1969

[Hof07] **P. Hofstedt, A. Wolf**, Einführung in die Constraint-Programmierung. Grundlagen, Methoden, Sprachen, Anwendungen
Springer, 2007

[Hop94] **J. Hopcroft, J. Ullman**, Einführung in die Automatentheorie, Formale Sprachen und Komplexitätstheorie
Addison-Wesley, 3. Aufl., 1994

[Hür02] **D. Hürlimann**, Objektorientierte Modellierung von Infrastrukturelementen und Betriebsvorgängen im Eisenbahnwesen
Dissertation, Schriftenreihe des Instituts für Verkehrsplanung und Transportsysteme der ETH Zürich, Heft 125, 2002

[Hun93] **R. Hundt**, Automatische Konflikterkennung in der Zugüberwachung
Signal + Draht, 85 (1993), Heft 5, S. 140-142

[Hyo97] **M. Hyoudou et al.**, An expert system for train operation adjustment of Shinkansen
IFAC-TS 97, Proceedings of the 8th IFAC Symposium on Transportation Systems, 1221-1226, Chania, 16.-18.6.1997

[Jac92] **I. Jacobson, M. Christerson, P. Jonsson**, Object-Oriented Software Engineering – A Use Case Driven Approach
Addison-Wesley, 1992

[Jac03] **J. Jacobs**, Rechnerunterstütze Konfliktermittlung und Entscheidungsunterstützung bei der Disposition des Zuglaufs
Dissertation, Veröffentlichungen des Verkehrswissenschaftlichen Institutes der RWTH Aachen, Heft 61, 2003

[Jack99] **P. Jackson**, Introduction to Expert Systems
Addison Wesley Longman, 3e, 1999

[Joc99] **H. E. Jochim**, Verkehrswirtschaftliche Ermittlung von Qualitätsmaßstäben im Eisenbahnbetrieb
Dissertation, Veröffentlichungen des Verkehrswissenschaftlichen Institutes der RWTH Aachen, Heft 54, 1999

[Jov89] **D. Jovanovic**, Improving railroad on-time performance: Models, Algorithms and Applications
Dissertation, University of Pennsylvania, 1989

[Jov91] **D. Jovanovic, P. Harker**, Tactical Scheduling of Rail Operations: The SCAN I System
Transportation Science, Vol. 25, S. 46-64, 1991

[Kir83] **S. Kirkpatrick, C. D. Gelatt, M. P. Vecchi**, Optimization by Simulated Annealing Science, Vol 220, Number 4598, pages 671-680, 1983

[Kla94] **V. Klahn**, Die Simulation großer Eisenbahnnetze
Dissertation, Wissenschaftliche Arbeit Nr. 41, Institut für Verkehrswesen, Eisenbahnbau und -betrieb der Universität Hannover, 1994

[Kem97] **A. Kemper, A. Eickler**, Datenbanksysteme – Eine Einführung,
Oldenbourg, 1997

[Kos90] **M. Kosemund, K. Danowski, A. Schmidt**, Über erste Versuche zur Anwendung wissensbasierter Systeme zur Betriebsführung einer Eisenbahn
Wissenschaftliche Zeitschrift der HfV, 37, 4, 617-620, 1990

[Kra87] **E. Kraft**, A branch and bound procedure for optimal train dispatching, Journal of Transportation Research Forum
Vol. 28, S. 263-276, 1987

[Kra95] **D. Kraay, P. Harker**, Real-time scheduling of freight railroads
Transportation Research, 29B(3), S. 213-229, 1995

[Kro94] **A. Krone, H. Kiendl**, Automatic Generation of Positive and Negative Rules for Two-Way Fuzzy Controllers
Second EUFIT (European Congress on Fuzzy and intelligent Technologies) Aachen, S. 438-447, 1994

[Kun05] **W. Kunz**, Eisenbahnrecht
Loseblattsammlung, Nomos, 26. Auflage, 2010

[Lan60] **A. Land, A. Doig**, An Automatic Method for Solving Discrete Programming Problems
Econometrica, Vol.28 (1960), S. 497-520

[Lau10] **W. Laun**, Regelbasiertes Programmieren in der Bahntechnik
Signal + Draht (102), 4/2010, S. 12-17

[Leh05] **H. Lehmann**, Die Anfänge der Betriebsüberwachung und -leitung bei den deutschen Eisenbahnen
Eisenbahningenieur (56) 5/2005, S. 52-55

[Mai53] **D. Maier**, The theory of relational databases
Computer Science Press, 1953

[Man97] **R. Manthey, U. Griefahn**, Aktive Datenbanken
Folien zur Vorlesung an der Universität Bonn, WS 1997/98

[Man08] **R. Manthey**, Deduktive Datenbanken
Folien zur Vorlesung an der Universität Bonn, SS 2008

[Mar95] **U. Martin**, Verfahren zur Bewertung von Zug- und Rangierfahrten bei der Disposition
Dissertation, Schriftenreihe des Instituts für Verkehr, Eisenbahnwesen und Verkehrssicherung der TU Braunschweig, Heft 52, Braunschweig, 1995

[Mit96] **M. Mitchell**, An Introduction to Genetic Algorithms
MIT Press, 1996

[NeO08] Deliverable 2.3 des Projektes New Opera (New European Wish – Operating project for European rail network) der Europäischen Kommission (6. Rahmenprogramm)
2008

[OCL06] **Object Management Group (OMG)**, Object Constraint Language OMG Specification Version 2.0, May 2006

[Pac93] **J. Pachl**, Steuerlogik für Zuglenkanlagen zum Einsatz unter stochastischen Betriebsbedingungen
Dissertation, Schriftenreihe des Instituts für Verkehr, Eisenbahnwesen und Verkehrssicherung der TU Braunschweig, Heft 49, 1993

[Pac04] **J. Pachl**, Systemtechnik des Schienenverkehrs – Bahnbetrieb planen, steuern und sichern
Teubner, 4. Aufl., 2004

[PBefG90] Personenbeförderungsgesetz vom 8.8.1990
BGBl. I S. 1690
zuletzt geändert am 29.7.2009 (BGBl. I S. 2258)

[Pep06] **P. Pepper, P. Hofstedt**, Funktionale Programmierung: Sprachdesign und Programmiertechnik
Springer, 2006

[Pet86] **E. Petersen, A. Taylor, C. Martland**, An introduction to computer assisted train dispatch
Journal of Advanced Transportation, 20(1), S. 63-72, 1986

[Pfe95] **S. Pferdmenges, H. Schaefer**, Automatische Konflikterkennung und wissensbasierte Konfliktlösung in der Streckendisposition
Signal + Draht 87 5/1995, S. 174-177

[Poh02] **P. Pohlmann**, Betriebszentralen- und ESTW-Planungen der DB Netz AG
Eisenbahntechnische Rundschau, 51 (2002), H. 7/8, S.410-416

[Rau95] **O. Rauh,** ERM^{ded} – eine Erweiterung des Entity-Relationship-Modells zur Modellierung deduktiver Informationssysteme
Informatik Forschung und Entwicklung, 10: 128 – 138, Springer Verlag, 1995

[Rau98] **R. H. Rauschenberg**, Potentiale für die Verringerung der externen Effekte des Verkehrssektors durch einen dezentralisierten und automatisierten Gütertransport der Bahn
Diss., Johann Wolfgang von Goethe Universität Frankfurt am Main, 1998

[Rei90] **W. Reisig**, Petrinetze – Eine Einführung
Springer, 1990

[Ros92] **M. Rosebrock**, Automatisierung und Dezentralisierung des Güterverkehrs der Bahn
Verlag Peter Lang, Frankfurt am Main, 1992

[Rum93] **J. Rumbaugh et. al.**, Objektorientiertes Modellieren und Entwerfen
Hanser, 1993

[Sak06] **C. Sakowitz, E. Wendler**, Optimising train priorites to support the regulation of train services with the assistance of active and deductive databases
In: Computers in Railways (Comprail) X (2006), S. 489-499.

[Sau83] **R. Sauder, W. Westerman**, Computer-aided train dispatching: decision support through optimisation
Interfaces 13(6), S. 24-37, 1983

[Scha05] **T. Schaer et al.**, DisKon – Laborversion eines flexiblen, modularen und automatischen Dispositionsassistenzsystems
Eisenbahntechnische Rundschau 54 (2005) H. 12, S. 809-821

[Scho06] **S. Scholz & T. Albrecht**, ALFa – a software tool for optimal scheduling of demand oriented train services
Computers in Railways X, S. 541-550, 2006

[Schw97] **W. Schwanhäußer**, Verfahren zur Steuerung und Sicherung eines fahrplangebundenen Verkehrssystems
Offenlegungsschrift DE 197 26 542 A, Anmeldenummer 3010244 vom 7.5.1997

[SQL08] SQL:2008 ISO/IEC 9075:2008
2008, siehe www.iso.org

[Sti05] **Stinnes Logistics**, Classic, Quality, Prime: die Einzelwagenprodukte im Schienengüterverkehr
Stand: April 2005

[Szp73] **B. Szpigel**, Optimal train scheduling on a single track railway
in: M. Ross (Ed.), OR'72, North-Holland, Amsterdam, 1973, S. 343-351

[Tak06] **R. Takagi et al.**, Optimal train control at a junction in the main line rail network using a new object-oriented signalling system model
Computers in Railways X, S. 479-488, 2006

[Tur36] **A. Turing**, On Computable Numbers, With an Application to the Entscheidungsproblem
Proceedings of the London Mathematical Society, Series 2, Volume 42, 1936; reprinted in M. David (ed.), The Undecidable, Hewlett, NY: Raven Press, 1965

[Ull88] **J. Ullman**, Principles of Database and Knowledge Base Systems
Computer Science Press, 1988/89

[UIC04] UIC-Merkblatt 406, Kapazität
2004

[UIC09] UIC-Merkblatt 450-2, Bewertung der Qualität auf dem Netz, in Bezug auf den Eisenbahnbetrieb, zur Durchführung von Qualitätsanalysen – Codierung von Verspätungen und Verfahren für die Zuordnung von Verspätungsursachen
5. Ausgabe, 2009

[UML04] **Object Management Group**, UML Spezifikation Version 2.0
2004, siehe www.omg.org

[Wal91] **K. Walther**, Maßnahmenreagibler Modal-Split für den städtischen Personenverkehr – Theoretische Grundlagen und praktische Anwendung
Veröffentlichungen des Verkehrswissenschaftlichen Institutes der RWTH Aachen, Heft 45, 1991

[Web02] **W. Weber**, Dispositive Regeln für die Reihenfolge von Zügen
BahnPraxis 11/2002, S. 116-117

[Weg05] **S. Wegele**, Echtzeitoptimierung für die Disposition im Schienenverkehr
Dissertation, Technische Universität Braunschweig, Institut für Verkehrssicherheit und Automatisierungstechnik, 2005

[Wen83] **D. Wende**, Fahrdynamik
Transpress VEB Verlag für Verkehrswesen, Berlin, 1983

[Wen99] **E. Wendler**, Analytische Berechnung der planmäßigen Wartezeiten bei asynchroner Fahrplankonstruktion
Dissertation, Veröffentlichungen des Verkehrswissenschaftlichen Instituts der RWTH Aachen, Heft 55, 1999

[Wen05] **E. Wendler, N. Nießen**, Infrastrukturmodellierung für die strategische Netzplanung
In: Proc. 20. Verkehrswissenschaftliche Tage, Dresden (2005), Session 4b.

[Wes71] **H. Westermayer**, Programmierlogik, Programmablaufpläne
Oldenbourg, 1971

[Wun75] **W. Wunderlich, E. Kockelkorn**, Bessere Disposition durch schnellere Information
Eisenbahntechnische Rundschau 24 (1975) H. 7/8, S. 283-294